DOCUMENT RESUME

ED 239 118 CE 038 060

TITLE	Unified Technical Concepts. Application Modules Volume II.
INSTITUTION	Center for Occupational Research and Development, Inc., Waco, Tex.
PUB DATE	83
NOTE	676p.; For related documents, see ED 190 746-761, CE 037 874-875, and CE 038 059.
AVAILABLE FROM	Center for Occupational Research and Development, 601C Lake Air Drive, Waco, TX 76710 ($12.00).
PUB TYPE	Guides - Classroom Use - Materials (For Learner) (051)
EDRS PRICE	MF04 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS	Behavioral Objectives; Energy; *Force; *Kinetics; Laboratory Experiments; Learning Modules; Motion; *Physics; Postsecondary Education; *Power Technology; *Radiation; Secondary Education; *Technical Education
IDENTIFIERS	Time Constants; *Transducers; Vibrations (Noise); Wave Generators

ABSTRACT

Unified Technical Concepts (UTC) is a modular system for teaching applied physics in two-year postsecondary technician programs. This UTC laboratory textbook, the second of two volumes, consists of 45 learning modules dealing with basic concepts of physics. Addressed in the individual chapters of the guide are the following topics: force transformers, energy converters, transducers, vibrations and waves, time constants, and radiation. Each chapter contains a series of modules that include an introductory statement, objectives, instructional text, a summary, a laboratory experiment, and a self-check. (MN)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED239118

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

D M Hull

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

CE038060

© 1983 by the Center for Occupational Research and Development

All rights reserved. No part of this book
may be reproduced in any form or by any
means without permission in writing from
the publisher.

The Center for Occupational Research and Development
601 C Lake Air Drive
Waco, Texas 76710

Library of Congress Catalog Card Number: 83-72030

PREFACE

for APPLICATION MODULES VOLUME II

The textbooks entitled UNIFIED TECHNICAL CONCEPTS and UTC APPLICATION MODULES form an integral set of teaching materials for the presentation of UTC Physics. UNIFIED TECHNICAL CONCEPTS is the "classroom text" designed for teaching the basic technical concepts in physics and demonstrating their wide application to the world of the technician. UTC APPLICATION MODULES--in two volumes--is the "laboratory text" which complements the subject matter found in the classroom text and provides the technician with valuable practice and application of physics concepts in the laboratory. The rationale for UTC Physics, the identification of its essential parts, and its importance in the education and training of modern technicians are covered in the Preface of UNIFIED TECHNICAL CONCEPTS--and are not repeated here. Suffice it to say that a successful presentation of UTC Physics requires an integrated presentation of the unified technical concepts in the classroom and meaningful experiments and measurements with modern-day devices and instruments in the laboratory.

In combination, Volumes I and II of UTC APPLICATION MODULES contain a total of 111 modules--carefully selected from the original 170 application modules (laboratory exercises) developed for UTC Physics. Volume I contains 56 application modules that illustrate and apply the concepts of FORCE, WORK, RATE, MOMENTUM, RESISTANCE, POWER and ENERGY to practical, real-world problems. Volume II contains 55 application modules that illustrate and apply the concepts of FORCE TRANSFORMERS, ENERGY CONVERTORS, TRANSDUCERS, VIBRATIONS AND WAVES, TIME CONSTANTS, and RADIATION to similar problems. In both Volumes I and II--as an aid to organization and presentation--all of the application modules associated with one particular technical concept are grouped together in a major section of each text and are identified as such at the beginning of each section dealing with that particular concept. For example, in Volume II, under the technical concept entitled TRANSDUCERS, there are 13 application modules included to illustrate and apply the concept of a "transducer" to meaningful laboratory experiments and measurements in four important easily recognized energy systems--mechanical, fluidal, electrical and thermal.

The notation we have adopted for identifying and keeping track of the many application modules deserves a word of explanation. We have used a three-symbol code--such as 1M1, 2F2, 9E1, or 10T4. The number of the left--varying from 1 to 13--refers to the particular concept which the application module supports. In this numbering system, FORCE → 1, WORK → 2, RATE → 3, etc. The letter in the middle refers to the particular energy system--mechanical (M), fluidal (F), electrical (E) and thermal (T). The number on the right refers to the particular module in the series written for a given concept (left number) and given energy system (middle letter). Thus the code 10T4 signifies that this application module is the fourth in a series written in support of concept 10 (TRANSDUCERS) dealing with a thermal (T) energy system. As the table of contents indicates, 10T4 is entitled "Thermocouples"--clearly a topic involved with thermal transducers.

For the most part the application modules contained in Volumes I and II of UTC APPLICATION MODULES were chosen to provide the student with important hands-on practice and application of technical concepts to meaningful, modern devices and instruments--and provide, at the same time, a broad experience in the mechanical, electrical, fluidal and thermal energy systems found in modern electromechanical technologies. A small group of application modules contained herein were chosen because their content appropriately supports and further strengthens the thirteen principal concepts covered in the classroom text. In some instances this group of modules introduces new ideas not found in the classroom text. These application modules, referred to appropriately as CONCEPT SUPPORT MODULES, are identified with an asterisk (*) in the table of contents for each of the two volumes and also in the list of application modules that precedes each of the major thirteen concepts, e.g., FORCE, WORK, etc.

As alluded to earlier, 170 application modules were developed for UTC Physics originally and all are available. The 111 modules we have included in the two volumes here represent our choice of the best of the 170. However, it has been our experience with different schools and faculties that some of the 59 not included in the two volumes are at times desirable. Therefore, we have listed in an appendix at the end of each volume, a title and brief description of each module not contained in either of these two volumes. One or all of these may be ordered from CORD as needed.

Finally, as teachers talking with teachers, we would strongly recommend that the implementation of UTC Physics not include more than 90 application modules in any full-year program--3 quarters or 2 semesters. In general, we advise that one begin with 60-80 modules in the first year. Then after the logistics in the laboratory have been worked out and teaching experience has been acquired, an increase to 90 (or beyond) can be considered. For planning and scheduling purposes, one should count on an average of about two hours per laboratory application module.

It goes without saying that a successful presentation of UTC Physics requires integrated teaching (and learning) in the classroom and the laboratory. We consider that portion of student time spent in the laboratory on meaningful, hands-on, technician-oriented exercises to be of vital importance. Overall success of UTC Physics requires a dedicated, enthusiastic teacher. It is not an easy course to teach, but it is exciting--and the rewards are great.

CONTENTS

		Page
Preface		iii
Acknowledgments		vii
Concept 8-0	FORCE TRANSFORMERS	1
8M1	Levers	2
8M3	Drive Systems	12
8M4	Gear Trains	22
8F1	Hydraulic Press	34
8E1	Electrical Transformers	44
Concept 9-0	ENERGY CONVERTORS	59
9F1	Turbines	60
9F2	Fluid Pumps	74
9F4	Fans and Blowers	89
9F5	Vacuum Pumps	101
9E1	Electric Generators	117
9E2	Alternators	129
9E3	Electric Motors	143
9E4	Solenoids and Relays	158
9E5*	Photovoltaic Materials	168
9T2	Thermoelectric Generators	180
9T3	Solar Collectors	197
Concept 10-0	TRANSDUCERS	212
10M1	Strain Gages	213
10M2	Accelerometers	225
10F1	Bourdon Tubes	232
10F3	Differential Pressure Transducers	241
10F4	Fluid-Flow Measuring Devices	250
10F5	Vacuum Gages	258
10E1	Meter Movements	275
10T2	Bimetallic Strips	288
10T3	Temperature Measuring Devices	299
10T4	Thermocouples	312
10T5	Thermistors	328
10T6	Pyrometers	340

* Identifies concept support modules that introduce new physics.

Page

Concept 11-0 VIBRATIONS AND WAVES 352
11M3 Vibration Isolation 353
11M4* Resonance 363
11F1* Cavity Vibrations and Standing Waves 374
11F4 Doppler Effect 382
11E1* Alternating Voltage and Current 398
11E2* Phase Relationships in AC Circuits 409
11E3 Oscilloscope 425
11E4* Rectifiers 437

Concept 12-0 TIME CONSTANTS 446
12F1 Rate of Emptying Tanks 447
12E3* RLC Circuits 461
12T1 Response Time of Thermocouples 471
12T2 Time Constants of Heating and Cooling 486

Concept 13-0 RADIATION 496
13L1* Inverse Square Law — Spreading of Light 497
13L2* The Optical Spectrum 509
13L3 Light Sources 522
13L4* Lasers 531
13L5 Stroboscopes — Light Pulses 540
13L6* Optical Filters — Absorption of Light 547
13L7* Specular and Diffuse Reflection of Light 558
13L8* Image Formation With Mirrors 570
13L9* Prisms — Refraction and Dispersion of Light 583
13L10* Lenses — Focusing and Spreading of Light 594
13L11* Interference — Testing Flat Surfaces 606
13L12* Diffraction Grating — Measuring the Wavelength of Light 615
13L13* Polarization — Glare Reduction 625
13L15 Spectroscopy — Light Analysis 638
13N3* Radioactivity 649
Appendix 664

* Identifies concept support modules that introduce new physics.

ACKNOWLEDGMENTS

Unified Technical Concepts (UTC)--Physics for Technicians is a major innovation in technical education for the 1980s. UTC Physics is innovative because it provides a broad, interdisciplinary base for teaching technical principles, combined with the practical illustration of these principles in the work clothes of modern industrial equipment.

The evolution of this delivery system is due largely to the contribution of several outstanding technical educators: First, to Dr. Maurice W. Roney, for his conceptualization of unified physics while developing an electromechanical technology curriculum in the late 1960s. Next, to Dr. Walter J. Brooking for his encouragement to persist in the development and dissemination of the UTC curriculum. Finally, to Dr. Leno S. Pedrotti the chief editor, technical consultant and designer of the instructional format. Significant contributions to UTC were also made by CORD staff members Johnny Jones and Carl Lansford, and by project editor Frank L. Pedrotti, S.J.

Daniel M. Hull, UTC Project Director
President, Center for Occupational
Research and Development
June 1983

FORCE TRANSFORMERS

APPLICATION MODULES

		Page
8M1	Levers	2
8M3	Drive Systems	12
8M4	Gear Trains	22
8F1	Hydraulic Press	34
8E1	Electrical Transformers	44

8M1

LEVERS

The lever is one of the simplest and most common mechanical force transformers. Depending upon its configuration, the lever may be used to amplify either force or displacement, or to change their directions.

In this module, the student will investigate the lever as a force transformer. In the laboratory, the student will measure and calculate the mechanical advantages of several levers.

OBJECTIVES

Upon completion of this module, the student should be able to:

- Calculate the fourth quality, given three of the following ones:
  - Load force.
  - Applied force.
  - Lever arm of load force.
  - Lever arm of applied force.
- Given the arm lengths of a lever, calculate its ideal mechanical advantage (M_i).
- Given the load force and applied force of a lever, calculate the actual mechanical advantage (M_a).
- Determine actual mechanical advantage and efficiency of a lever, given its dimensions.
- Sketch and identify the three classes of levers; give two examples each of how they are used.
- Given the appropriate equipment, determine the actual mechanical advantage and the efficiency of a lever.

DISCUSSION

Figure 1 shows the forces applied to a class 1 lever. In this figure, assume that the lever is a rigid rod with no mass and that all forces are applied at right angles to the lever. The force F_1 is the load force and has a lever arm of ℓ_1: F_2 is the applied force and has a lever arm of ℓ_2. The force F_p is the force applied by the pivot, or fulcrum.

In most applications of the lever, the system is very near the equilibrium condition; that is, the sum of both the forces and the torques on the lever is zero:

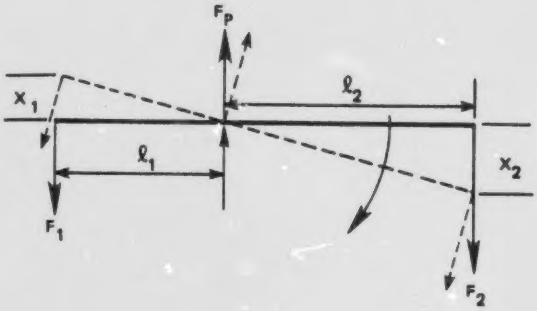


FIG. 1

$$\Sigma F = F_1 + F_2 - F_p = 0$$

$$(\Sigma T)_{pivot} = F_1\ell_1 - F_2\ell_2 = 0$$

where

ΣF = sum of forces on lever.
ΣT = sum of torques on lever at pivot.
F_1 = load force of lever arm ℓ_1.
F_2 = applied force of lever arm ℓ_2.
F_p = force applied by pivot, or fulcrum.

This equation represents the condition of the lever when there is no motion or when there is motion of the lever at a constant rate.

If the pivot of an ideal lever is assumed to be fixed, the sum of forces must remain zero; however, the sum of torques can be other than zero. If the lever is initially at rest in the equilibrium condition, and if F_2 is increased, F_p increases; and the sum of the forces remains zero. Increased torque $F_2\ell_2$ is not offset, however, by an increase in $F_1\ell_1$ (because neither F_1 nor ℓ_1 changed); and a clockwise angular acceleration takes place. If F_1 is increased later, an opposing angular acceleration is created; and the lever rotates in the opposite direction. If F_2 moves through a vertical distance of x_2, F_1 then moves through a vertical distance of x_1 in the opposite direction. Equation 1 gives the relationships between all quantities for a lever.

$$\frac{F_2}{F_1} = \frac{\ell_1}{\ell_2} = \frac{x_1}{x_2} \qquad (1)$$

where

x_1 = vertical displacement (distance) of load force in one direction.

x_2 = vertical displacement (distance) of applied force in opposite direction to x_1.

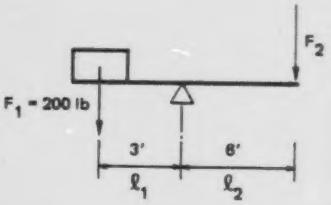
Example A shows use of Equation 1 in solving a problem.

EXAMPLE A: LEVER ARM CALCULATION

Given: The lever shown is used to raise a 200-lb load.

Find: (a) Value of F_2 for equilibrium and (b) height of the load if F_2 moves downward 0.5 feet.

Solution:



F_2
F_1 = 200 lb
3'
6'
ℓ_1
ℓ_2

(a) $\frac{F_2}{F_1} = \frac{\ell_1}{\ell_2}$ or $F_2 = \frac{F_1 \ell_1}{\ell_2}$

$$F_2 = \frac{(200\ \text{lb})(3\ \text{ft})}{(6\ \text{ft})}$$

$$= \frac{600\ \text{lb-ft}}{6\ \text{ft}}$$

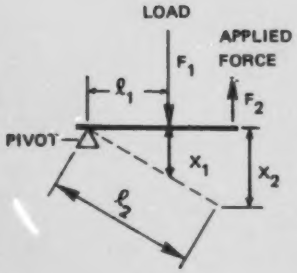
$$F_2 = 100\ \text{lb}$$

(b) $\frac{x_1}{x_2} = \frac{\ell_1}{\ell_2}$ or $x_1 = \frac{x_2 \ell_1}{\ell_2}$

Therefore,

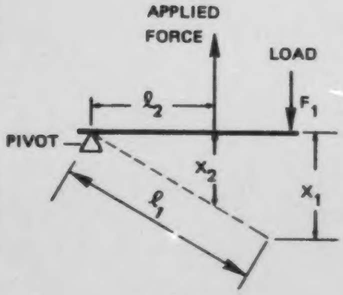
$$x_1 = \frac{(0.5)(3)}{6}$$

$$x_1 = \frac{1.5}{6} = 0.25\ \text{ft}$$



LOAD
APPLIED FORCE
F_1
ℓ_1
F_2
PIVOT
x_1
x_2
ℓ_2

a. CLASS II



APPLIED FORCE
LOAD
ℓ_2
F_1
PIVOT
x_2
x_1
ℓ_1

b. CLASS III

FIG. 2

Figure 1 shows a class I lever, in which the fulcrum always is located between the load force and the applied force. Figure 2 shows two additional classes of levers.

In a class II lever, the fulcrum is at one end of the lever; the applied force is at the other end; and the load is between the two. In this type lever, the load force is always greater than the applied force.

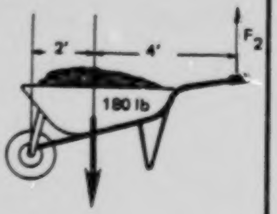
In a class III lever, positions of the load force and applied force are exchanged. The applied force always is greater than the load force.

Equation 1 applies to all classes of levers. Examples B and C show the use of Equation 1 in solving additional problems.

EXAMPLE B: WHEELBARROW LOAD CALCULATION

Given: A loaded wheelbarrow weighs 180 lb and its center of gravity is 2 ft from the wheel axle. The lifting force is applied 4 feet farther from the wheel axle.

Find: (a) Applied force and (b) how far the load rises when the applied force rises 15 inches.



Solution:

(a) $$F_2 = \frac{F_1\ell_1}{\ell_2} = \frac{(180\ \text{lb})(2\ \cancel{\text{ft}})}{6\ \cancel{\text{ft}}}$$

$$F_2 = 60\ \text{lb}$$

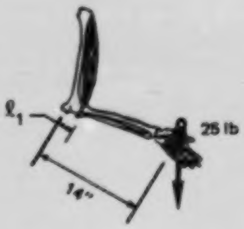
(b) $$x_1 = \frac{x_2\ell_2}{\ell_2} = \frac{(15\ \text{in})(2\ \text{ft})}{(6\ \text{ft})} = \frac{(15\ \text{in})(24\ \text{in})}{(72\ \text{in})}$$

$$x_1 = 5\ \text{in}$$

EXAMPLE C: HUMAN ANATOMY STRENGTH CALCULATION

Given: The length of a person's arm from elbow to palm is 14 inches. When the biceps are shortened by 1/4 inch, the 25-lb weight in the hand is raised by 1 3/4".

Find: (a) Tension in the person's muscle and (b) lever arm of the applied force.



Solution:

(a) $F_2 = \frac{F_1 x_1}{x_2} = \frac{(25\ \text{lb})(1.75\ \cancel{\text{in}})}{(0.25\ \cancel{\text{in}})}$

$F_2 = 175\ \text{lb}$

(b) $\ell_1 = \frac{\ell_2 x_2}{x_1} = \frac{(14\ \text{in})(0.25\ \cancel{\text{in}})}{(1.75\ \cancel{\text{in}})}$

$\ell_1 = 2\ \text{in}$

Mechanical Advantage for an ideal lever is given by Equation 2.

$$M_i = \frac{\ell_2}{\ell_1} \tag{2}$$

where

M_i = ideal mechanical advantage.
ℓ_1 = lever arm of load force.
ℓ_2 = lever arm of applied force.

In the ideal case, the assumption often is made that the lever itself has no mass and that there is no friction at the pivot point. Because this assumption never really is the case, we usually find that the actual mechanical advantage usually is found to be different from the ideal. The actual mechanical advantage is defined by Equation 3.

$$M_a = \frac{F_1}{F_2} \tag{3}$$

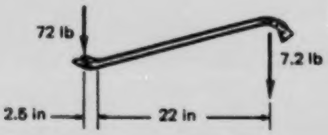
where

M_a = actual mechanical advantage.
F_1 = load force.
F_2 = applied force.

Example D shows the use of Equations 2 and 3 in solving a problem.

EXAMPLE D: LOAD LIFTING USING A CROWBAR

Given: A crowbar used to lift a 72-lb load vertically with an applied force of 7.2 lb. The respective lever arms are 2.5 in and 22 in.



Find: Efficiency of the crowbar.

Solution:

$$M_i = \frac{\ell_2}{\ell_1} = \frac{22\ \text{in}}{2.5\ \text{in}} = 8.8$$

$$M_a = \frac{F_1}{F_2} = \frac{72\ \text{lb}}{7.2\ \text{lb}} = 10$$

$$\%\ \text{efficiency} = \frac{M_a}{M_i} \times 100 = \left(\frac{10}{8.8}\right)(100) = 114\%$$

Notice that, in this example, the efficiency is greater than 100%, the reason being that the weight of the heavy metal bar has been neglected. The downward gravitational force acting at the center of gravity of the crowbar actually assists the applied lifting force.

APPLICATION OF LEVERS

Every conscious person capable of movement uses one or more of the lever classes each day. Examples of this simple and most common of all force transformers are abundant in contemporary life. The crowbar is one of the purest examples of the class I lever. More obscure examples of the class I lever are scissors, pliers, and the rocker arm assembly of an internal combustion engine. The paper cutter of figure 3 shows one of the clearest examples of class II levers. Class III levers are abundantly represented in the human body and visible in machines such as the boom crane. In figure 3 a few of the abundant examples are shown with reference to the applied force (F_A) and the load force (F_L).

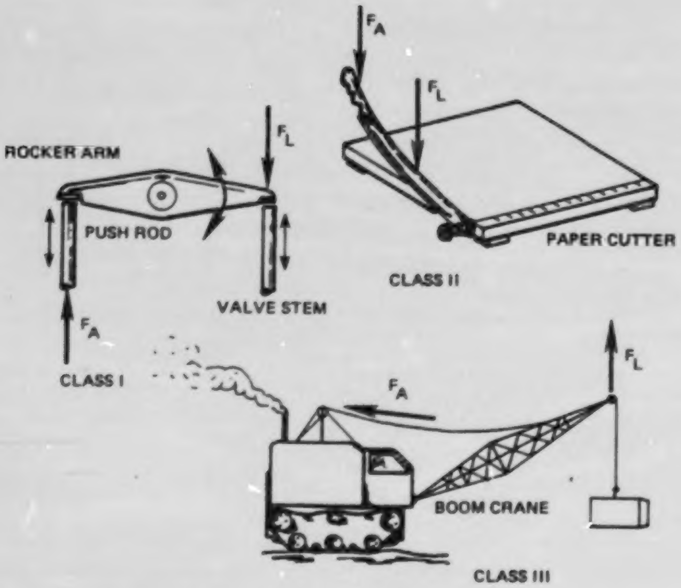

ROCKER ARM
PUSH ROD
VALVE STEM
F_L
F_A
CLASS I
F_A
F_L
PAPER CUTTER
CLASS II
F_L
F_A
BOOM CRANE
CLASS III

FIG. 3 APPLICATION OF LEVERS

SUMMARY

The lever is a common, simple force transformer that can amplify force or displacement and can change their directions. Class I levers change direction of force and amplify it. Class II levers amplify force without changing its direction. Class III levers amplify displacement without changing direction of force.

MECHANICAL ADVANTAGE OF LEVERS

CLASS I LEVER

MATERIALS

Eight-inch piece of 2 inch x 4 inch wood.
Two-foot piece of 1 inch x 8 inch wood.
Spring balance with 50-lb capacity.
Crowbar.

PROCEDURE

Note: All loads can consist of weights or sandbags of same weight as measured with bathroom scales.

1. Position wooden block, board, crowbar and spring balance as shown in Figure 4.

2. Have a classmate stand in center of board; apply tension to spring balance until board is level. Record values of ℓ_1, ℓ_2 and F_2 in Data Table.

3. Weigh classmate; record one half weight of classmate as F_1 in Data Table.

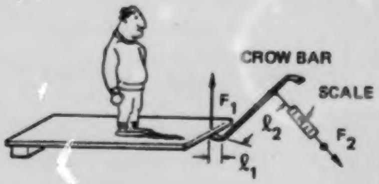


FIG. 4

CLASS II LEVER

MATERIALS

Bathroom scales.
Set of English weights or sandbags.
1/2-inch diameter steel bar 4 ft long.

PROCEDURE

1. Hang a weight of 40 to 60 lb at approximate center of steel bar, as shown in Figure 5. Have classmate stand on bathroom scales, and hold one end of bar while other end rests on a table.

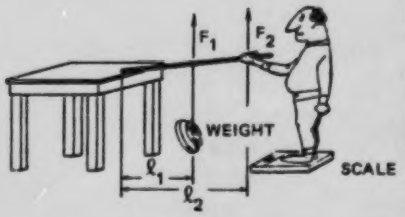


FIG. 5

2. Measure and record ℓ_1, ℓ_2 and reading from bathroom scales.
3. Record F_1 as weight hung on bar and F_2 as difference between bathroom scale reading and classmate's weight.

CLASS III LEVER

MATERIALS

Set of English weights or sandbags.
Spring balance with 50-lb capacity.
1/2-inch diameter steel bar 4 ft long.

PROCEDURE

1. Hang a 10-lb weight from end of steel bar, as shown in Figure 6. Place other end of bar under top of a heavy table; support bar about one third of way from pivot with spring balance.

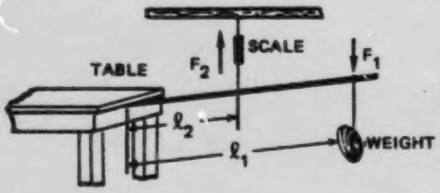


FIG. 6

2. Measure ℓ_1, ℓ_2, F_1 and F_2; record them in Data Table. Perform all necessary calculations; complete Data Table.
3. Explain any differences in % efficiency between the three classes of levers tested.

DATA TABLE							
Class	ℓ_1 (in)	ℓ_2 (in)	F_1 (lb)	F_2 (lb)	M_i	M_a	% efficiency
I							
II							
III							

SELF CHECK

1. A 95-lb boy sits on a seesaw 6 feet from the pivot. How far from the pivot must a 85-lb girl sit to balance him?

2. Draw sketches to identify the three classes of levers; list two examples of each class.

3. In a nutcracker (class II lever), a force of 32 lb is applied 6.5 inch from the fulcrum. What is the force on a nut 1.25 inch from the fulcrum?

4. In a machine for throwing clay targets, a force of 210 lb is applied to a lever 4 inch from the fulcrum. What force is experienced by the clay target 24 inch from the fulcrum?

5. A class II lever is used as an input force transformer to drive a hydraulic jack. An applied force of 50 lb produces a downward force of 600 lb on the hydraulic cylinder. The lever arms are 1.5 inch and 19 inch. Find (a) ideal mechanical advantage, (b) actual mechanical advantage and (c) percentage efficiency of the lever.

6. A type I lever is used to lift a 400 lb block. If the lever is 90% efficient (under all conditions), what is the minimum applied force that the lever must be able to withstand if it is to be used to lift the block? The load lever arm is 1 ft and the lever arm of the applied force is 4 ft.

8M3

DRIVE SYSTEMS

In rotational mechanical systems, the forcelike quantity is torque. Torque transformers called "drive systems" are employed when changes in torque and angular velocity are desired.

In this module, the student will investigate gear systems, belt drives and disk drives as examples of torque transformers. In the laboratory, the student will examine the disk drive of a turntable.

OBJECTIVES

Upon completion of this module, the student should be able to:

- Calculate mechanical advantage in the following systems:
  Gear drive system, given number of teeth on each gear.
  Belt drive system, given radius of each wheel.
  Disk drive system, given radius of each disk.
- Given mechanical advantage of a drive system, input torque and angular velocity, calculate output torque and angular velocity.
- Given the appropriate equipment, determine mechanical advantage, input and output angular velocities and torques of the disk drive of a record player turntable, given necessary data to solve appropriate equations.

DISCUSSION

Figure 1 shows a torque transformer called a belt drive. During some time t, a point on the belt moves from position A to position B. If there is no slipping, a point on wheel 1 moves through arc A_1B_1 (equal to distance $\overline{AB}$) and a point on wheel 2 moves through an equal arc A_2B_2 (also equal to distance $\overline{AB}$). The angular velocities ω of the two wheels are as follows:

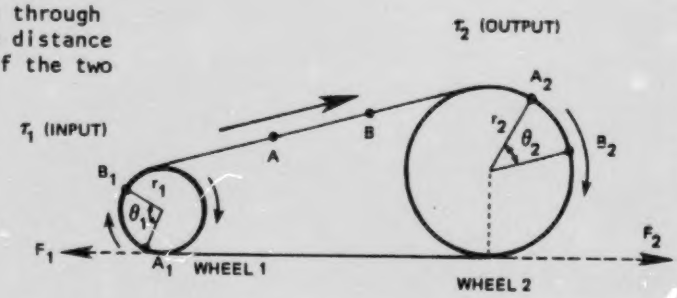


FIG. 1

$$\text{since } \theta_1 = \frac{\text{arc } A_1B_1}{r_1}$$

$$\omega_1 = \frac{\theta_1}{t} = \frac{\text{arc } A_1B_1}{r_1 t}$$

and

$$\text{since } \theta_2 = \frac{\text{arc } A_2B_2}{r_2}$$

$$\omega_2 = \frac{\theta_2}{t} = \frac{\text{arc } A_2B_2}{r_2 t}.$$

Since arc A_1B_1 equals arc A_2B_2, dividing one equation by the other results in Equation 1.

$$\frac{\omega_1}{\omega_2} = \frac{r_2}{r_1} \tag{1}$$

where
ω_1 = angular velocity of wheel 1.
ω_2 = angular velocity of wheel 2.
r_1 = radius of wheel 1.
r_2 = radius of wheel 2.

If F_1 and F_2 are tangential forces at the rims of the wheels. torques acting on the two wheels are:

$$\tau_1 = F_1 r_1$$

and

$$\tau_2 = F_2 r_2.$$

F_1 and F_2 each equals the tension in the belt, and they must be equal to each other. After dividing, Equation 2 becomes:

$$\frac{\tau_1}{\tau_2} = \frac{r_1}{r_2} \qquad (2)$$

where
τ_1 = torque of wheel 1.
τ_2 = torque of wheel 2.

Mechanical advantage of the transformer is defined as ratio of the output torque to input torque. Equations 1 and 2 are combined to obtain Equation 3.

$$M = \frac{\tau_2}{\tau_1} = \frac{r_2}{r_1} = \frac{\omega_1}{\omega_2} \qquad (3)$$

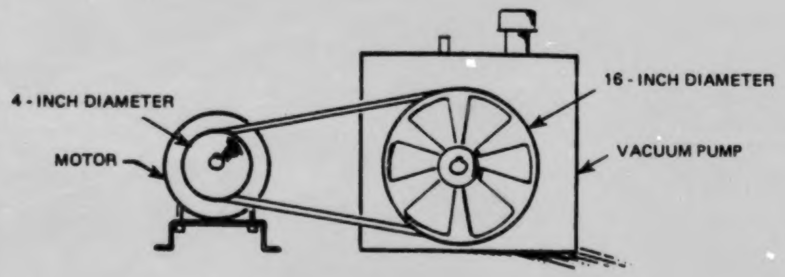
where
M = mechanical advantage.

An increase in torque results in a corresponding decrease in angular velocity. Example A shows use of Equation 3 in solving a problem.

EXAMPLE A: BELT DRIVE SYSTEM

Given: A belt drive system vacuum pump has a motor wheel with a 4-inch diameter and a 16-inch diameter load wheel. The motor torque is 4 ft·lb, and the motor turns at 480 rpm.

Find:
(a) Mechanical advantage of the drive system.
(b) Torque supplied to the load.
(c) Rotational rate of the pump.



Solution: From Equation 3:

(a) $M = \frac{r_2}{r_1}$

$$r_1 = \frac{4}{2} = 2 \text{ in}$$

$$r_2 = \frac{16}{2} = 8 \text{ in}$$

$$M = \frac{8 \text{ in}}{2 \text{ in}} = 4$$

(b) $M = \frac{\tau_2}{\tau_1}$

$$\tau_2 = M\ \tau_1$$

$$\therefore \tau_2 = (4)\ (4 \text{ ft}\cdot\text{lb}) = 16 \text{ ft}\cdot\text{lb}$$

(c) $M = \frac{\omega_1}{\omega_2}$

$$\omega_2 = \frac{\omega_1}{M}$$

$$= \frac{480 \text{ rpm}}{4}$$

$$\therefore \omega_2 = 120 \text{ rpm}$$

(d) Another solution (from Equation 3) for c would be:

$$\frac{r_2}{r_1} = \frac{\omega_1}{\omega_2}$$

$$\therefore \omega_2 = \omega_1 \frac{r_1}{r_2} = 480 \left(\frac{\overset{1}{\cancel{2}}}{\underset{4}{\cancel{8}}}\right)$$

$$\omega_2 = 120 \text{ rpm}$$

Figure 2 shows a disk drive system in which one disk drives another by friction contact. A careful comparison of this diagram with Figure 1 shows that the only difference is the direction of rotation of the second wheel. The disk drive in this figure reverses direction of rotation; a belt drive does not, unless the belt is crossed. All previous equations apply to both disk and belt drives.

Figure 3 shows a gear drive system identical to the disk drive system, except that planetary teeth are cut on the circumference of each disk to create a stronger, more efficient drive, with the number of teeth on each gear proportional to the ra-

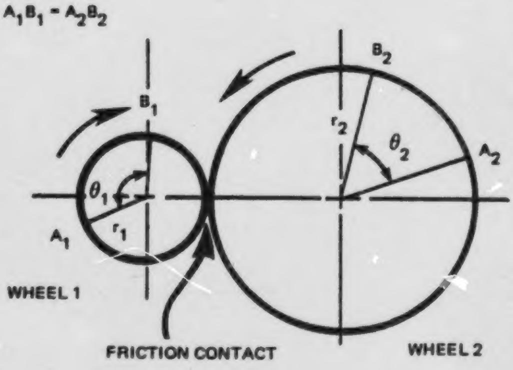


FIG. 2

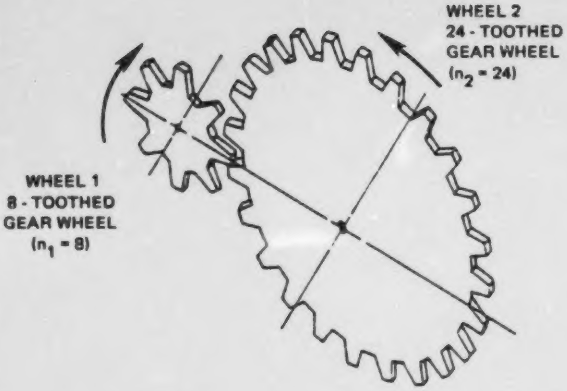

WHEEL 2
24 - TOOTHED
GEAR WHEEL
($n_2 = 24$)
WHEEL 1
8 - TOOTHED
GEAR WHEEL
($n_1 = 8$)

FIG. 3

dius of the gear. Equation 4 is rewritten to obtain Equation 4.

$$M = \frac{\tau_2}{\tau_1} = \frac{n_2}{n_1} = \frac{\omega_1}{\omega_2} \qquad (4)$$

where

n_1 = number of teeth on gear 1.

n_2 = number of teeth on gear 2.

Example B shows use of Equation 4 in solving a problem.

EXAMPLE B: GEAR DRIVE IN CLOCKS

Given: In an alarm clock, a gear with 55 teeth drives a gear with 7 teeth. The first gear has a rotational rate of 3 revolutions per day.

Find: (a) Mechanical advantage of the drive and (b) rotational rate of the second gear.

Solution:

(a) $M = \frac{n_2}{n_1}$

$= \frac{7}{55}$

$M = 0.1273$

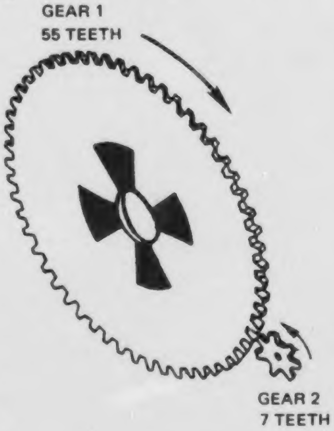
(b) $M = \frac{\omega_1}{\omega_2}$

from which

$\omega_2 = \frac{\omega_1}{M}$

$= \frac{3 \text{ r/day}}{0.1273}$

$\omega_2 = 23.57 \text{ r/day}$


GEAR 1
55 TEETH
GEAR 2
7 TEETH

APPLICATIONS OF DRIVE SYSTEMS

Since rotational motion is a convenient method of transferring mechanical power from one point to another, drive systems are normally used with electric motors and internal combustion engines. Gear drives are common in clocks, engine lathes and automobile transmissions. Belt drives often are employed with electric motors in pumps, blowers and tape recorders. In an automobile, belts drive the fan, water pump and power brakes. Disk drives are common in systems requiring accurate constant angular velocity and low torque such as turntables of stereo record players.

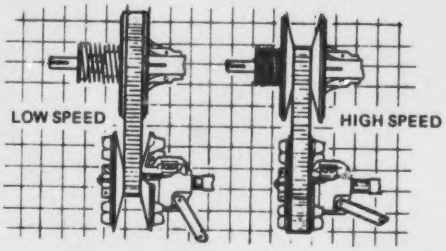


FIG. 4

Figure 4 shows a variable-diameter pulley system of a drill press. The wide V-belt rides on two cone, or variable-diameter, pulleys. Each pulley is composed of two conical sections whose separation may be varied. The pulley with the cones closest together has the largest effective diameter.

Figure 5 shows the disk drive system of a record player turntable. The motor shaft with a radius r_1 drives a disk of radius r_2. This disk drives the larger turntable (radius r_3) from the inside. Example C is a problem concerning this figure.

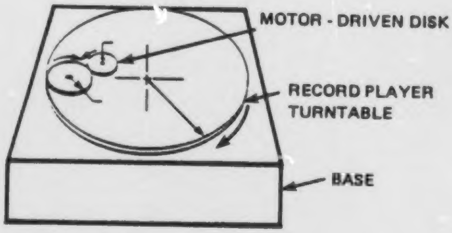


FIG. 5

EXAMPLE C: MULTI DISK DRIVE

Given: Figure 5; the motor rotates at 3600 rpm and $r_1 = 0.785$ mm, r_2 1.85 cm, $r_3 = 8.5$ cm.

Find: Rotational rate of turntable and mechanical advantage of the drive system.

Solution:

$$M\ (r_1 \text{ to } r_2) = \frac{r_2}{r_1}$$

$$= \frac{1.85\text{ cm}}{0.0785\text{ cm}}$$

$$= 23.57$$

$$M\ (r_2 \text{ to } r_3) = \frac{r_3}{r_2}$$

$$= \frac{8.5 \text{ cm}}{1.85 \text{ cm}}$$

$$= 4.60$$

$$\text{System } M\ (r_1 \text{ to } r_3)$$

$$= M\ (r_1 \text{ to } r_2) \times M\ (r_2 \text{ to } r_3)$$

$$= (23.57)\ (4.60)$$

$$\therefore \text{system } M = 108$$

$$\text{Now system } M = \frac{\omega_1}{\omega_3}$$

$$\therefore\ \omega_3 = \frac{\omega_1}{M} = \frac{3600}{108} = 33\text{-}1/3 \text{ rpm}$$

Notice that this is the standard rpm for the long-playing record.

SUMMARY

Drive systems transform torque into a desired angular velocity or an angular velocity into a desired torque. The primary types of drive systems are gear, belt and disk drives. They normally are used with electric motors and/or internal combustion engines.

DISK DRIVE OF A TURNTABLE

MATERIALS

Phonograph turntable.
Micrometer calipers.
Rule (12 in).
Map tack.
Light spring balance (16 oz).

PROCEDURE

1. Remove turntable from its shaft; examine drive mechanism (Figure 6). Note: Turntable is driven from inside by a rubber disk, which is driven by a shaft turned by a motor. The motor shaft has four sections, each of different diameter. If smallest diameter is used, table turns at 16 rpm. If largest diameter is used, table turns at a rate of 78 rpm. The other two sections are for 33-1/3 and 45 rpm. Measure diameter of rubber disk and diameter of 33-1/3 rpm section of motor shaft. Record this data in Data Table.

2. Measure and record inner diameter of turntable.

3. Reassemble turntable.

4. Press tack into turntable surface at a point near rim. Attach spring balance as shown in Figure 7; position it so its line of action is tangent to turntable. Set turntable angular speed for 33-1/3 rpm, and switch on. Record scale reading and value of ℓ in Data Table.

5. Complete calculations in Data Table.

6. Repeat above steps for a second case with an angular speed of 45 rpm.

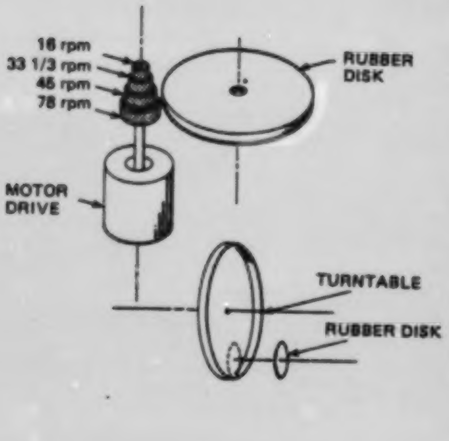

16 rpm
33 1/3 rpm
45 rpm
78 rpm
RUBBER DISK
MOTOR DRIVE
TURNTABLE
RUBBER DISK

FIG. 6

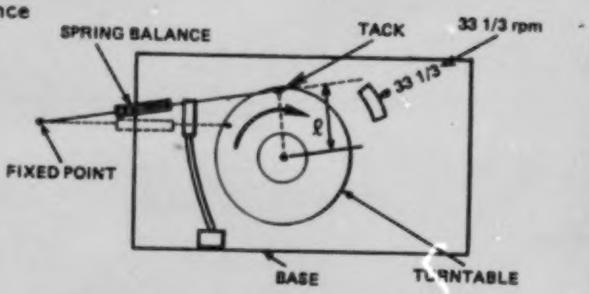

SPRING BALANCE
TACK
33 1/3 rpm
33 1/3
ℓ
FIXED POINT
BASE
TURNTABLE

FIG. 7

7. What is the relationship of angular velocity and mechanical advantage to torque?

DATA TABLE

Case 1: ω_3 = 33.3 rpm.

Motor shaft diameter: d_1 = ______ inches.

Disk diamter: d_2 = ______ inches.

Turntable inner diameter: d_3 = ______ inches.

Force on spring balance: F = ______ ounces.

Lever arm of spring balance: ℓ = ______ inches.

Mechanical advantage: $M = \left(\frac{d_3}{d_2}\right)\left(\frac{d_2}{d_1}\right)$

$= \frac{d_3}{d_1} =$ ______ inches ÷ ______ inches

$\therefore M =$ ______.

Input angular velocity: $\omega_1 = \omega_3 M$

ω_1 = 33.3 rpm x ______

ω_1 = ______ rpm.

Output torque: $\tau_3 = F\ell$

= ______ ounces x ______ inches

τ_3 = ______ ounces x inches.

Input torque: $\tau_1 = \frac{\tau_3}{M} = \tau_3 \div M$

τ_1 = ounces x inches ÷ ______

τ_1 = ______ ounces x inches.

Case 2: ω_3 = 45.0 rpm.

d_1 = ______ inches.
d_2 = ______ inches.
d_3 = ______ inches.

These are the same as measured in Case 1.

F = ______ ounces.

ℓ = ______ inches (same as measured in Case 1).

M = ______.

ω_1 = ______ rpm.

τ_3 = ______ ounces x inches.

τ_1 = ______ ounces x inches.

SELF CHECK

1. A gear with 18 teeth drives a gear with 56 teeth. What is the system mechanical advantage?

2. In a belt drive system for a band saw, the motor has a pulley with a diameter of 2-12 inches, and the load has a pulley with a diameter of 14 inches. What is the system mechanical advantage?

3. The disk drive of a turntable has three disks with diameters of 5.75 in, 1.15 in and 0.1065 in. What is the system mechanical advantage?

4. The input torque of the motor in the turntable in problem 3 is 0.42 oz x in and the turntable is rotating at 33-1/3 rpm. Find output torque and angular velocity of the motor.

5. State reasons for differences in ω_1 and τ_1 between Cases 1 and 2 in Data Table.

6. In a gear drive system the large wheel has a diameter of 30 cm and a total of 40 teeth. How many teeth does the small gear have if it has a diameter of 6 cm? What is the mechanical advantage of the system?

8M4

GEAR TRAINS

One of the more important types of force transformers is the gear train. No other force transformation system is used more widely as a means of producing and applying mechanical energy. Gear drive systems commonly are employed to change torque, angular velocity and direction of rotation in rotational mechanical systems. This module discusses the design, characteristics and applications of a variety of common gear drive systems. In the laborabory, the student will examine the gear drive of a metal cutting lathe.

OBJECTIVES

Upon completion of this module, the student should be able to:

- List three characteristics of rotational motion that can be changed by gear trains.
- Given diagrams of the following gear trains, identify each:

  Simple.
  Compound.
  Reverted.
  Planetary.
  Differential.
- Given the number of gears and teeth on each gear in the following gear trains, determine the mechanical advantage and direction of rotation of the output shaft:

  Simple.
  Compound.
  Reverted.
- Draw and label diagrams of a planetary gear train, showing the direction of rotation of each gear under the following conditions:

Input on sun gear, output on cage, ring gear locked.
Input on sun gear, output on ring gear, cage locked.
Input on ring gear, output on cage, sun gear locked.
Input on cage, output on ring gear, sun gear locked.

- Draw and label a diagram of a simple automobile differential gear train. Show the rotation of each gear during cornering.

- State one application of each of the following gear trains:

  Compound.
  Reverted.
  Planetary.
  Differential.

- Given the gear train of a lathe with a Norton gear box, determine the gear ratio between the works spindle and feed shaft for two different feed speeds.

THE SIMPLE GEAR TRAIN

Figure 1 shows a simple gear drive that consists of two gears. This system transforms rotational motion of the input shaft to output rotational motion with a different angular velocity, torque (forcelike quantity) and direction of rotation. Equation 1 gives the ideal mechanical advantage, also called the gear ratio, of such a gear train.

$$M = \frac{\tau_2}{\tau_1} = \frac{N_2}{N_1} = \frac{\omega_1}{\omega_2} \qquad (1)$$

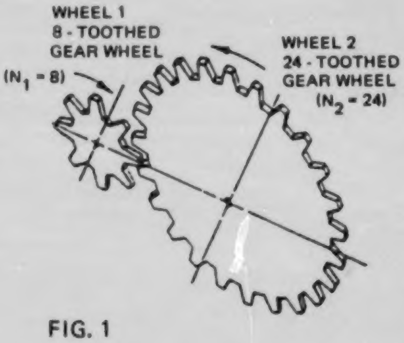


FIG. 1

where

M = mechanical advantage.
τ_1 = torque on gear 1.
τ_2 = torque on gear 2.
N_1 = number of teeth on gear 1.
N_2 = number of teeth on gear 2.
ω_1 = angular velocity of gear 1.
ω_2 = angular velocity of gear 2.

Example A shows the use of Equation 1 in solving a problem.

EXAMPLE A: MECHANICAL ADVANTAGE OF A SIMPLE GEAR TRAIN

Given: In Figure 1, wheel 1 has a rotational rate of 2.4 rad/s and delivers a torque of 0.4 N•m.

Find:
(a) Mechanical advantage.
(b) Rotational rate of wheel 2.
(c) Output torque of wheel 2.

Solution: $M = \frac{\tau_2}{\tau_1} = \frac{N_2}{N_1} = \frac{\omega_1}{\omega_2}$

(a) $M = \frac{N_2}{N_1}$

$= \frac{24}{8}$

$M = 3$

(b) $\omega_2 = \frac{\omega_1}{M}$

$= \frac{2.4 \text{ rad/s}}{3}$

$= \omega_2 = 0.8$ rad/s

(c) $\tau_2 = M\tau_1$

$= (3)(0.4 \text{ N}\cdot\text{m})$

$\tau_2 = 1.2$ N•m

A simple gear train is one in which the centers of all gears are fixed, and there is only one gear per shaft. Figure 2 is another example of such a gear train. The direction of rotation of the output is the same as that of the input if the train contains an odd number of gears and is opposite in direction for an even number of gears. The mechanical advantage of such a train is determined by applying Equation 1 for each pair of interlocking gears, as shown in Example B.

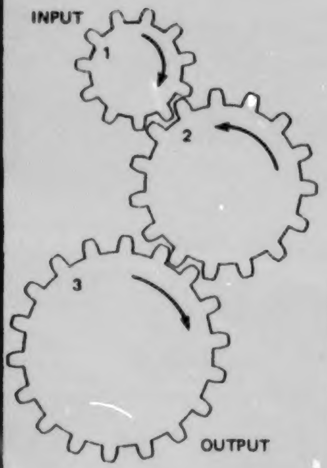


FIG. 2

EXAMPLE B: MECHANICAL ADVANTAGE OF A SIMPLE TRAIN WITH MORE THAN TWO GEARS

Given: A four-gear train has gears with gears having the following numbers of teeth: $N_1 = 16$, $N_2 = 8$, $N_3 = 13$, $N_4 = 17$.

Find: Mechanical Advantage.

Solution:

$$M\ (1\ \text{to}\ 2) = \frac{N_2}{N_1} = \frac{8}{16}$$

$$M\ (2\ \text{to}\ 3) = \frac{N_3}{N_2} = \frac{13}{8}$$

$$M\ (3\ \text{to}\ 4) = \frac{N_4}{N_3} = \frac{17}{13}$$

$$M\ (\text{total}) = M\ (1\ \text{to}\ 2)\ M\ (2\ \text{to}\ 3)\ M\ (3\ \text{to}\ 4)$$

$$= \left(\frac{\cancel{N_2}}{N_1}\right)\left(\frac{\cancel{N_3}}{\cancel{N_2}}\right)\left(\frac{N_4}{\cancel{N_3}}\right)$$

$$= \frac{N_4}{N_1}$$

$$M = \frac{17}{16}$$

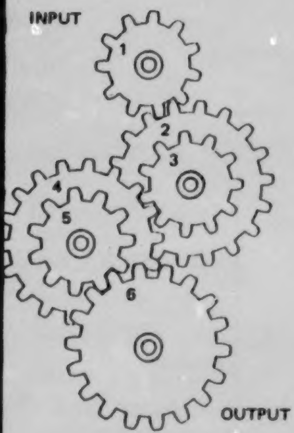


FIG. 3

Example B indicates that in a simple train, the mechanical advantage depends only upon the number of teeth on the input gear and the number on the output gear. Intermediate gears do not effect mechanical advantage.

THE COMPOUND GEAR TRAIN

Figure 3 shows a compound gear train in which some shafts carry two gears that are rigidly connected to rotate at the same rate and with the same torque. The same technique used to calculate the mechanical advantage of a simple gear train can be applied to the compound train, as shown in Example C.

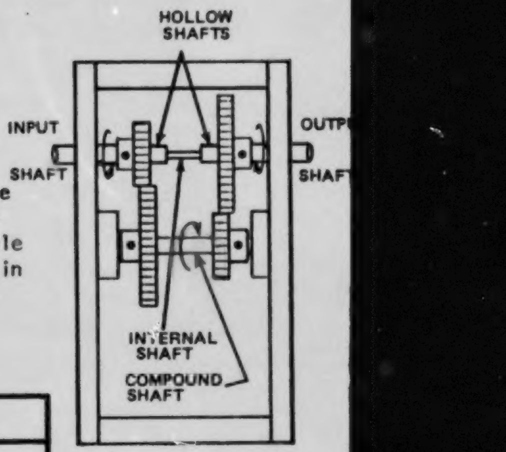

HOLLOW SHAFTS
INPUT SHAFT
OUTP
SHAF
INTERNAL SHAFT
COMPOUND SHAFT

FIG. 4

EXAMPLE C: MECHANICAL ADVANTAGE OF A COMPOUND GEAR TRAIN

Given: Gear train in Figure 3, $N_1 = 11$, $N_2 = 18$, $N_3 = 11$, $N_4 = 18$, $N_5 = 11$, $N_6 = 18$.

Find: Mechanical Advantage.

Solution: The overall mechanical advantage is the product of the mechanical advantage for all gear pairs.

$$M\ (1 \text{ to } 2) = \frac{N_2}{N_1} = \frac{18}{11}$$

$M\ (2 \text{ to } 3) = 1$ because gears are rigidly connected.

$$M\ (3 \text{ to } 4) = \frac{N_4}{N_3} = \frac{18}{11}$$

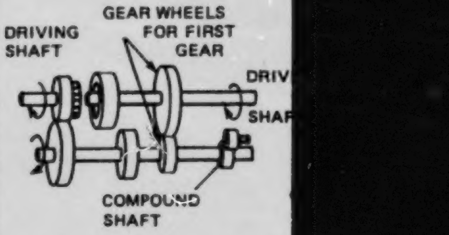
$$M\ (4 \text{ to } 5) = 1$$

$$M\ (5 \text{ to } 6) = \frac{N_6}{N_5} = \frac{18}{11}$$

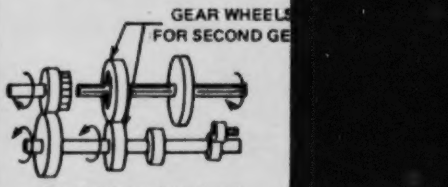
$$M \text{ total} = \left(\frac{N_2}{N_1}\right)\left(\frac{N_4}{N_3}\right)\left(\frac{N_6}{N_5}\right) = \frac{N_2 N_4 N_6}{N_1 N_3 N_5}$$

$$= \frac{(18)(18)(18)}{(11)(11)(11)}$$

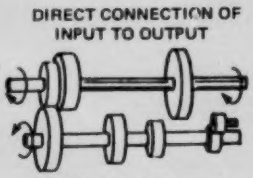
$$= 4.38$$


GEAR WHEELS FOR FIRST GEAR
DRIVING SHAFT
DRIV
SHAF
COMPOUND SHAFT

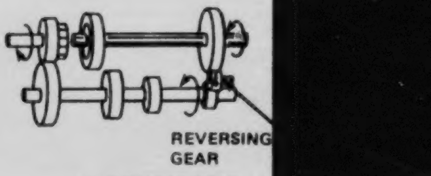
(a) FIRST GEAR (LOW)


GEAR WHEEL
FOR SECOND GE

(b) SECOND GEAR (MID)


DIRECT CONNECTION OF INPUT TO OUTPUT

(c) THIRD GEAR (HIGH)


REVERSING GEAR

(d) REVERSE GEAR

FIG. 5

The compound train in Figure 3 is composed of three simple trains — gear 1 drives gear 2, gear 3 drives gear 4, and gear 5 drives gear 6. The mechanical advantage of the compound train is the product of the mechanical advantages of the simple train that it contains.

Compound gear trains are used in mechanical clocks and in drive trains for lathes, power tools, and automobiles. Figure 4 shows a special class of compound train, called a reverted train, in which the input and output shafts are coaxial. This design affords ease of changing gear ratios in a compact space. Its most familiar application is in the automobile transmission in Figure 5.

First gear provides the greatest mechanical advantage to give a high torque and low rotational rate for starting and low-speed operation. Second gear provides a lower mechanical advantage for intermediate speeds. In third gear, the input and output shafts are connected directly to give a mechanical advantage of one for high speed operation. In reverse, an additional gear called an idler is added to reverse the direction of rotation of the output shaft.

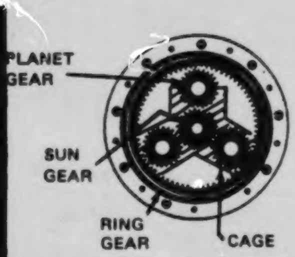


FIG. 6

THE PLANETARY GEAR TRAIN

Figure 6 shows a planetary gear train in which the centers of rotation of some gears are not fixed. This device consists of a central sun gear, one or more planetary gears attached to a carrier called a "cage," and an outer ring gear. The sun gear, cage, and ring gear can be connected to the input or output by concentric shafts. Planetary gears are used as force transformers in bicycles and automatic transmissions for automobiles.

Figure 7 shows how a planetary gear train is used in an automobile. The connections are as follows:

- Low gear: Sun gear — input.
  Ring gear — locked.
  Cage — output.

  This arrangement gives the greatest mechanical advantage and the lowest speed.

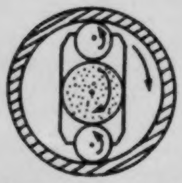
- Mid gear: Sun gear — locked.
  Ring gear — input.
  Cage — output.

  This configuration gives a lower mechanical advantage and a greater speed.

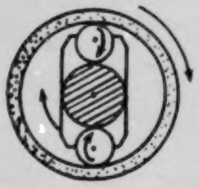
- High gear: Sun gear — locked.
  Ring gear — output.
  Cage — input.

  This connection gives the least mechanical advantage and the greatest speed.

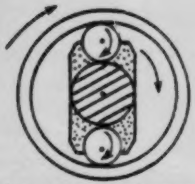
- Reverse gear: Sun gear — input.
  Ring gear — output.
  Cage — locked.



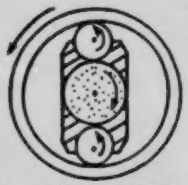
(a) FIRST GEAR (LOW)



(b) SECOND GEAR (MID)



(c) THIRD GEAR (HIGH)



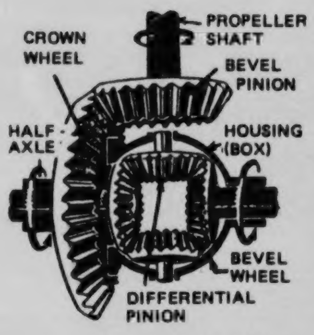
(d) REVERSE GEAR

FIG. 7

This arrangement reverses the direction of rotation of the output shaft.

The differential gear train in Figure 8 is a special application of the planetary train used to produce two outputs with equal torque, but differing rotational rates.

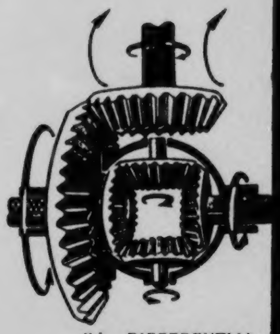
The example shown is the differential train of an automobile. The crown gear is rigidly attached to the housing. Each half axle is attached to a level gear. During straight travel (Figure 8a) the crown gear, housing and both half-axles all rotate at the same rate, and the differential pinions do not rotate with respect to the housing or other gears. During cornering (Figure 8b) one wheel must rotate faster than the other. In this case the pinion gears rotate as shown allowing the two half-axles to rotate at different rates but with equal torques.



(a) EQUAL OUTPUTS

SUMMARY

A simple gear drive consists of two gears that transform rotational motion of the input shaft to output rotational motion having a different angular velocity, torque (forcelike quantity) and direction of rotation. A compound gear train is composed of two or more simple trains; the mechanical advantage (M) of such a train is the product of the trains that it contains. The planetary gear is used for many special purposes because of its small size and unique qualities; for example, it is used in automobile differentials, where it produces two outputs with equal torque but different rotational rates.



(b) DIFFERENTIAL OUTPUTS (CORNERING)

FIG. 8

LABORATORY

DETERMINING GEAR RATIOS

MATERIALS

Lathe with Norton gear box (or other suitable gear train).
Tools to remove covers of gear box.
Pencil.

PROCEDURE

1. Disconnect electrical power of lathe and disengage lathe motor.

2. Remove covers to expose gear drive system that connects spindle shaft of lathe to feed shaft (Figure 9).

3. Count number of teeth on each gear and record numbers in data section.

4. Engage movable lever and gear for greatest feed speed as shown in Figure 9.

5. Rotate spindle shaft by hand in direction indicated in Figure 9. Observe direction of rotation of each gear in train and record by drawing appropriate arrows in Figure 9.

6. Calculate mechanical advantage between spindle shaft and feed shaft.

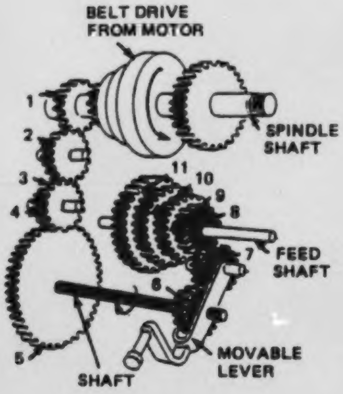


FIG. 9

7. Make small index mark on spindle shaft and feed shaft with pencil.

8. Rotate spindle shaft to produce one complete rotation of feed shaft. Count and record rotations of spindle shaft necessary. Estimate fractions of a rotation from final position of index mark on spindle shaft.

9. Repeat procedure for three different feed speeds.

10. Replace all cover plates.

DATA TABLE					
Gear	Teeth	Gear	Teeth	Gear	Teeth
1		5		9	
2		6		10	
3		7		11	
4		8			

Mechanical advantage (Refer to Figure 9):

$$M \text{ (1 to 8)} = \frac{N_3}{N_1} \times \frac{N_5}{N_4} \times \frac{N_8}{N_6}$$

M = ____________

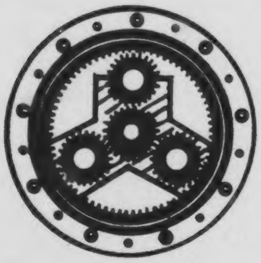
Number of rotations of spindle shaft for one rotation of feed shaft: N = __________

Output Gear	Mechanical Advantage (M)	Rotations (N)
8		
9		
10		
11		

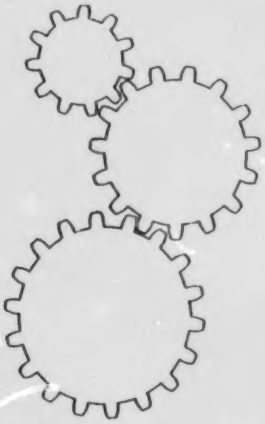
SELF CHECK

1. Gear trains may be used to change what three quantities?

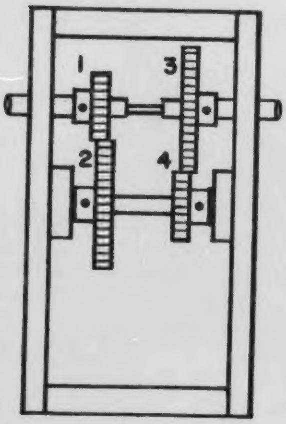
2. Identify each of the following types of gear trains.



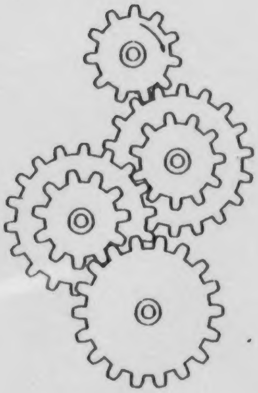
(a) ______________



(b) ______________



(c) ______________

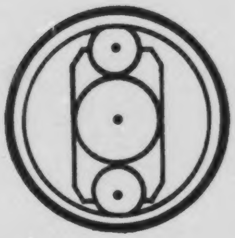


(d) ______________

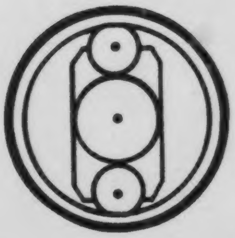
3. In the gear train shown in Problem 2c above, $N_1 = 19$, $N_2 = 26$, $N_3 = 18$, $N_4 = 27$. Find the mechanical advantage.

4. In the gear train shown in Problem 2d above, draw arrows indicating the direction of rotation of each gear.

5. Indicate the direction of rotation of the cage and each gear on the following diagrams. Assume input rotation is clockwise.

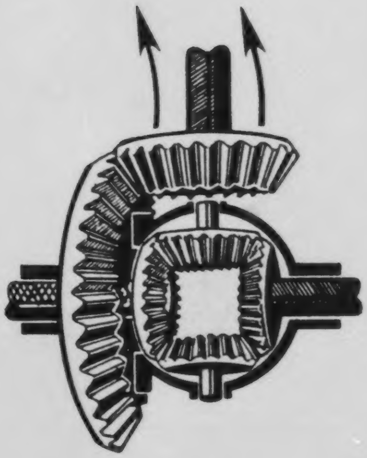


Input on sun gear.
Output on ring gear.
Cage locked.



Input on cage.
Output on ring gear.
Sun gear locked.

6. On the diagram at the right, indicate the direction of rotation of each gear as the automobile turns left.



7. List the type of gear train used in each of the following devices:

   a. Automatic transmission of automobile.
   b. Standard transmission of automobile.
   c. Mechanical clock.

8. A simple gear drive consists of two gears having 25 and 6 teeth respectively. If the gear with 6 teeth develops a torque 0.2 N•m, what torque does the other gear develop?

9. A simple gear train consists of 5 gears. The mechanical advantage of the system is 6. If the first gear in the sequence has 12 teeth, what can be said of the number of teeth on the other gears?

10. A compound gear train looks like the one shown in Figure 3 but the number of teeth on the gears is not the same as for the gears shown. If the mechanical advantage of the system is 6, and gears 1, 3, and 5 are identical, what is the ratio of the number of teeth on gear 1 to the number of teeth on gear 2?

11. A simple gear train like the one shown in Figure 1 is used in a clock. If a second hand is to be attached to gear 1 and a minute hand to gear 2, how many teeth should gear 2 have if gear 1 has 5 teeth?

HYDRAULIC PRESS

In one form or another, force transformers are used every day. They may not be identified as such, but lifting an automobile in the gas station repair area, pressing the brake pedal, or stamping out sheet metal parts in a factory, are all examples of simple force transformers in the fluid system. They all utilize the same physical principle — the hydraulic press — to gain a force advantage when the applied force must move through a greater distance than the load. This module introduces the hydraulic press. In the laboratory, the student will construct and operate a hydraulic-press simulator.

OBJECTIVES

Upon completion of this module, the student should be able to:

- Define:

    Closed hydraulic system.
    Hydraulic force transformation.
    Mechanical advantage.
    Hydraulic advantage.

- Using illustrations, describe Pascal's law.
- Given diameter of two pistons in a closed hydraulic system, calculate:

    Distance moved by one, given distance moved by the other.
    Lift of one, given the force applied to the other.

- List at least two examples of mechanisms that apply Pascal's law.
- Given the appropriate equipment, demonstrate the characteristics of a hydraulic press.

CLOSED HYDRAULIC SYSTEMS

A closed hydraulic system contains a rigid tank similar to the one in Figure 1, completely filled with a liquid and vented to several pistons as shown. There are no leaks of liquid through the piston sides, no air space at the top of the tank, and no liquid pumped in or out during operation.

If any of the pistons are moved in or out or up or down, one or more of the other pistons would have to move. This is because the liquid is not compressible. (Note that in an enclosed pneumatic system — air, not liquid — the air compresses easily, and the same laws do not apply.)

If piston A is moved inward, one piston, or a combination of the other pistons, would have to move outward, as no liquid or air could seep in, and the total volume of the system would remain the same.

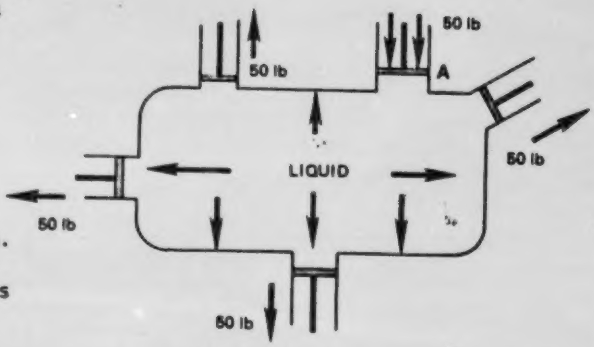


FIG. 1

PASCAL'S PRINCIPAL

According to Pascal's law, the pressure in an enclosed liquid tank is essentially the same everywhere within the tank. Furthermore, pressure is the same in all directions at any point; for if the pressure were different, there would be a flow of liquid until the pressure was equalized.

As a consequence of Pascal's law, two affects in the tank in Figure 1 can be observed. When an inward force of 50 lb is applied at piston A, it produces a pressure that is transmitted equally throughout the tank. A pressure measurement made anywhere in any direction would be the same. Furthermore, if all the pistons shown had the same area, the outward force of each would be 50 lb, the same as the force applied inwardly on piston A.

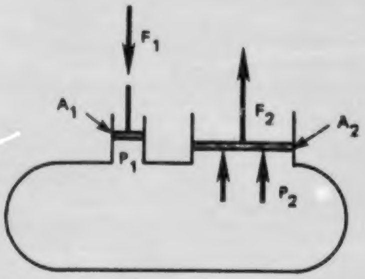


FIG. 2

FORCE MULTIPLICATION

Consider a tank with two pistons of different diameters, as a means of multiplying force.

In Figure 2, the force applied at piston A will produce a pressure in the tank. By Pascal's law, this pressure is transmitted equally throughout the fluid and, hence, is the same at piston A_2.

The basic formula for pressure is given by Equation 1.

$$p = \frac{F}{A} \tag{1}$$

where

p = pressure in pounds per square inch.
F = force applied in pounds.
A = area to which the force is applied in square inches.

Because pressure is the same everywhere in a closed hydraulic system, the pressure at piston 1 is the same as the pressure at piston 2, as indicated by Equation 2.

$$p_1 = p_2 \tag{2}$$

where

p_1 = pressure at piston 1.
p_2 = pressure at piston 2.

The pressure at piston one is shown by Equation 3.

$$p_1 = \frac{F_1}{A_1} \tag{3}$$

where

F_1 = force on piston 1.
A_1 = area of piston 1.

Similarly, Equation 4 gives the pressure at piston 2.

$$p_2 = \frac{F_2}{A_2} \tag{4}$$

where

F_2 = force at piston 2.
A_2 = area of piston 2.

Since $p_1 = p_2$, the force at piston 2 can be calculated by Equation 5.

$$\frac{F_1}{A_1} = \frac{F_2}{A_2} \qquad (5)$$

or

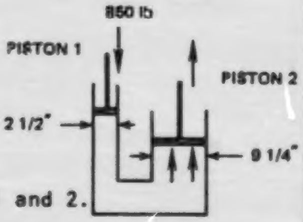
$$F_2 = \left(\frac{A_2}{A_1}\right) F_1$$

Thus the force at piston two can be greater than the applied force at piston one if $A_1 > A_2$. The ratio A_2/A_1 is called the mechanical advantage of the system. An example of the use of these equations is given in Example A.

EXAMPLE A: CALCULATION OF A FORCE MULTIPLICATION PROBLEM

Given: An 850-lb force is applied to piston 1, which has a diameter of 2.5". Piston 2 is 9.25" in diameter.

Find: Force applied on piston 2.



Solution: Using Equation 5:

$$\frac{F_1}{A_1} = \frac{F_2}{A_2}$$

Calculate the area of pistons 1 and 2.

$$\text{Piston 1 area} = \frac{\pi}{4} d^2$$

$$= \frac{\pi}{4} (2.5)^2$$

$\pi = 3.14$
$d = 2.5$ inches

$$= \left(\frac{3.14}{4}\right)(6.25)$$

$$A_1 = 4.91 \text{ square inches}$$

Similarly:

$$\text{Piston 2 area} = \frac{\pi}{4} (9.25)^2$$

$$= \left(\frac{3.14}{4}\right)(85.5)$$

$$A_2 = 67.2 \text{ square inches}$$

Rearranging Equation 5 gives:

$$F_2 = \left(\frac{A_2}{A_1}\right) F_1 = \left(\frac{67.2}{4.91}\right)(850)$$

$$F_2 = 11{,}633 \text{ lb}$$

Therefore, an 850-lb force on piston 1 will lift a force of 11,633 lb on piston 2.

EFFECT OF FORCE TRANSFORMATION ON DISTANCE

The analysis of work involved shows that work in equals work out. Work is equal to force applied times distance moved. If more force is put out, the distance traveled by the force will be less; therefore, except for frictional losses, the work output equals the work input. Force, not energy, is multiplied.

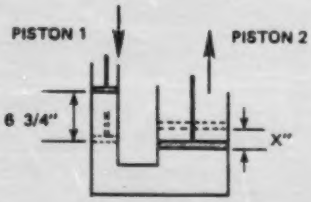
Consider again Example A, from the standpoint of distance traveled by each of the two pistons. See Example B.

EXAMPLE B: CALCULATION OF DISTANCE FORCES TRAVEL USING EXAMPLE A

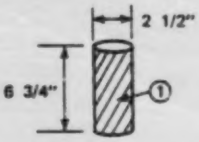
Given: Same hydraulic system as in Example A, and piston 1 moves 6 3/4".

Find: How far piston 2 moves.

Solution: (a) The solution to this problem may be found by assuming that the total volume of the system remains the same during piston movement; therefore —



Volume of liquid displaced by the movement of piston 1 = volume of liquid displaced by the movement of piston 2



$$\text{Volume of piston 1} = \left(\frac{\pi}{4} d_1^2\right)(6\ 3/4)$$

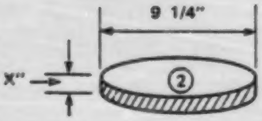
$$= (4.9 \text{ sq in})(6.75)$$

$$V = 33 \text{ cubic inches}$$

$$\text{Volume of piston 2} = \frac{\pi}{4} d_2^2(x)$$

$$V = 67.2\ (x)$$

$$\text{Volume 1} = \text{Volume 2}$$



Therefore,

$$33 = 67.2\ (x)$$

$$x = \frac{33}{67.2} = 0.5 \text{ in}$$

Note that in raising such a large weight of 11,633 lb, piston 2 only moves 1/2".

(b) A somewhat similar approach uses the work principle, as follows:

$$F_1x_1 = F_2x_2$$

$$x_2 = \left(\frac{F_1}{F_2}\right) x_1$$

$$= \left(\frac{850}{11,633}\right)(6.75)$$

$$x_2 = 0.5 \text{ in}$$

A hydraulic press is a further development of the hydraulic jack principle (Figure 3). If a frame is bolted to the hydraulic jack, considerable pressure can be exerted between the two platens to bend or squeeze metal into a variety of desired shapes.

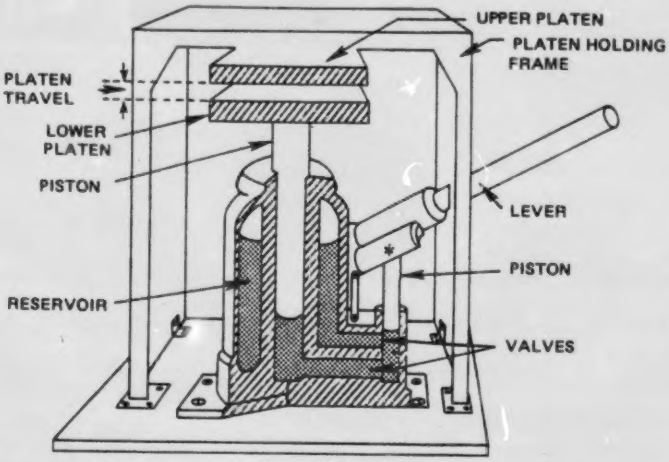

UPPER PLATEN
PLATEN HOLDING FRAME
PLATEN TRAVEL
LOWER PLATEN
PISTON
LEVER
PISTON
RESERVOIR
VALVES

FIG. 3

SUMMARY

The hydraulic press is a force transformer that transforms an applied force into a greater force at the output. The applied force must pass through a greater distance than the load. The force upon the liquid in a hydraulic press is the same in all directions for every square inch; however, when an output piston is larger (in cross-sectional area) than an input piston, the output piston exerts greater force, because more area is exposed to liquid pressure.

SIMULATION OF HYDRAULIC PRESS

MATERIALS

60-cm^3 disposable syringe — eccentric tip without needle.
12-cm^3 disposable syringe — eccentric tip without needle.
Eight 500-gm slotted weights.
Two each 200-g, 100-g and 50-g slotted weights.
One 5-gallon bucket or plastic container.
Flexible rubber or plastic tubing, clear 1/16 inch — inside diameter, 1/32 inch wall thickness (medical supply house).
Camps to secure hose to syringes.
Two ring stands and two clamps to support syringes in vertical position.
Vernier caliper.

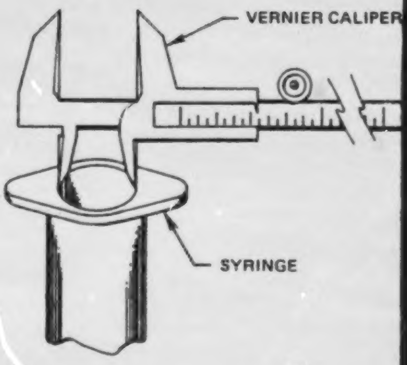

VERNIER CALIPER
SYRINGE

FIG. 4

PROCEDURE

Construct a simulated hydraulic press as follows:

1. Remove plungers from both 12-cm^3 and 60-cm^3 syringes. Measure inside diameter of cylindrical barrels with vernier calipers (Figure 4). Record these values in Data Table.

2. Place the two unassembled syringe barrels and plungers and a twelve-inch length of tubing into 5-gallon container full of water.

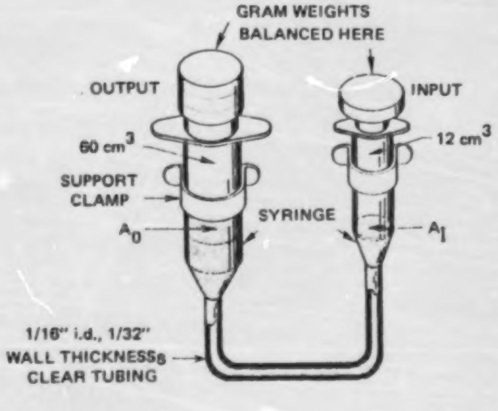

GRAM WEIGHTS BALANCED HERE
OUTPUT
INPUT
60 cm^3
12 cm^3
SUPPORT CLAMP
SYRINGE
A_0
A_I
1/16" i.d., 1/32" WALL THICKNESSs CLEAR TUBING

FIG. 5

3. Assembly of hydraulic press simulator in this experiment is made totally under water, so that no air is allowed to leak into closed systems. With 60-cm^3 syringe under water, push its plunger all the way to the botton to exhaust any air; fill with about 10 cm^3 of water. Leave it submerged at bottom of bucket; repeat with the 12-cm^3 plunger. Connect tubing, filled with water, to eccentric tips of both plungers, while holding both syringes

and tubing under water. Remove apparatus from water; mount syringes in a vertical position, making sure that they are secure and well supported (Figure 5).

4. Check operation of hydraulic system by pushing down manually first on one plunger, then on the other. If no air is in hydraulic system, plungers should move immediately and simultaneously.

5. System assembled (Figure 5) may be termed a hydraulic press simulator. The 60-cm^3 syringe represents output of simulator; and 12-cm^3 syringe represents input. Manually pushing input plunger requires a certain amount of force to make output plunger move. Force required is that necessary to overcome frictional resistance of hydraulic system; that is, drag of rubber plungers. To measure this breakway force, proceed as follows. With no weights on output plunger, carefully balance slotted weights on input plunger until weight is just sufficient to make both plungers begin to move and continue to move slowly. Record this weight (breakaway force) in Data Table. Remove weights; return fluid to input syringe (approximately 6-8 cm^3).

6. Carefully balance three 500-g slotted masses on output plunger, while manually preventing input plunger from being forced out of its tube. Place frictional force weight measured in No. 5 on input syringe; carefully stack just enough additional slotted weights on top of input plunger to start system moving slowly. This procedure should show clearly that a given force on input side causes a greater force (output weight) to be overcome. Record total force on input plunger in Data Table. Note: Under loaded condition of step 6, if weights are not carefully balanced and syringe barrels kept vertical, plungers rub against barrels, causing additional friction and adding to instability of apparatus. Error thus may be introduced into measurements associated with this experiment.

7. Repeat process outlined in steps 5 and 6 for two additional output weights; record all data in Data Table. Note: In this experiment masses are loaded on plungers. The fact that the following ratios are equal allows use of masses and weights interchangeably:

$$\frac{\text{Output mass load}}{\text{Input mass load}} = \frac{\text{Output force}}{\text{Input force}}$$

8. Complete the following exercises:

   (a) Calculate cross-sectional area of the output and input syringes with the following equation:

$$A = \frac{\pi}{4} d^2$$

Record results in Data Table.

(b) For each trial in the experiment, calculate difference between the input mass for that trial and the input mass from trial 1 (no output load). Record these differences in Data Table.

(c) Calculate ratio of the output force (mass) to the output area; record in Data Table.

(d) Calculate ratio of the true input force (i.e., total weight less frictional force) to the output force; record in the Data Table.

(e) Calculate ratio of the input cross-sectional area to the output cross-sectional area; compare the ratios obtained in exercises d and e. Record in the Data Table (see Equation 2).

DATA TABLE

Diameter of 60-cm^3 output syringe = ______ cm

Diameter of 12-cm^3 input syringe = ______ cm

Cross-sectional area of output (A_o) = ______ cm^2

Cross-sectional area of input (A_i) = ______ cm^2

Trial	Mass Unloaded On Output (g)	Mass Loaded On Input (g)	Input Mass Minus Frictional Mass
1	0		0
2			
3			
4			

F_o/A_o = _____ dynes/cm^2

F_i (true)/F_o = _____

A_i/A_o = _____

How closely does $\frac{F_i}{F_o} = \frac{A_o}{A_i}$?

SELF CHECK

1. Write a short description of the following terms, include any sketches to clarify presentation:

    Closed hydraulic system.
    Hydraulic force transformation.
    Mechanical advantage.
    Hydraulic pressure.

2. Define Pascal's law, using illustrations to describe the principles involved.

3. If a load of 1000 lb is moved through 21 inches on the input piston, how much load will it lift on the output piston if the input and output piston diameters are 1 inch and 12 inches respectively? How far will the 12-inch piston (output piston) move during this process?

4. Give two everyday examples of the use of Pascal's law.

5. In Figure 2, what would be the pressure on a small area at the bottom of the tank if $P_2 = 50\ \mathrm{lb/ft^2}$?

6. A certain thin-walled circular piston is made of a material that can withstand up to 1000 $\mathrm{lb/ft^2}$ before it fractures. What is the maximum force this piston can withstand if the piston's diameter is 3 inches?

7. Show that the two approaches to finding the distance the forces travel given in Example B are equivalent by using Equation 5.

8. The output piston of a hydraulic system does 100 joules of work by moving 20 cm. How far does the input piston move if the ratio of the area of the output piston to that of the input piston is 3 to 1?

8E1

ELECTRICAL TRANSFORMERS

The transformer, whose principal function is to increase or to decrease the voltage available from a source of alternating current, was one of the earliest electrical components, and is still one of the most widely-used ones.

This module introduces the operating principles of the transformer. Voltage, current, power relationships, and transformer power losses also will be investigated. In the laboratory, the student will construct an electrical transformer and will determine its characteristics.

OBJECTIVES

Upon completion of this module, the student should be able to:

- Given the turns ratio of the coils in a transformer, current (I_p) and voltage (V_p) input, calculate current (I_s) and voltage (V_s) output of the secondary coil.
- Given either voltages or currents in the primary and secondary coils of a transformer, calculate ratio of turns of the coils.
- Briefly discuss the three major sources of power loss in a transformer.
- Given the secondary resistance (R_s) of a transformer and the turns ratio of the coils, calculate the primary resistance (R_p).
- Given the appropriate equipment, determine the operating characteristics of an electrical transformer.

DISCUSSION

The principal function of a transformer is to increase or to decrease the voltage available from a course of alternating current, accomplished through magnetic induction. The changing of voltage in one wire loop will induce a voltage in a nearby loop. This induced voltage is due to the changing magnetic field that joins the two loops.

A simple transformer, as shown in Figure 1, consists of three principal parts: one, a primary coil connected to a source of alternating current; two, a secondary coil attached to a load (light bulb, appliance, etc.); and, three, a soft iron core.

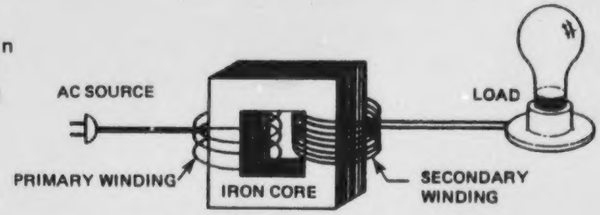


FIG. 1

The a.c. voltage source of the primary winding generates an alternating magnetic flux in the core that induces an a.c. voltage in the secondary winding. The magnitude of voltage in the secondary is dependent upon the number of turns in the two coils. Schematically, this simple transformer could be drawn as shown in Figure 2.

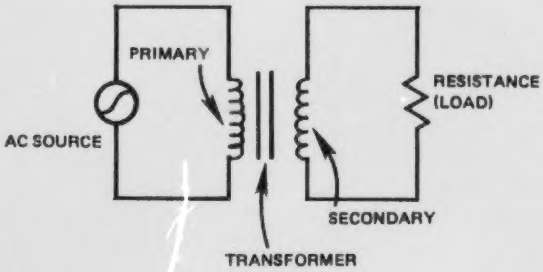


FIG. 2

CLASS OF TRANSFORMERS

Depending on whether they increase or decrease the voltage applied to them, there are essentially two classes of transformers — one that increases voltage, called a "step-up transformer," and one that decreases voltage, called a "step-down transformer." How a transformer increases or decreases voltage is explained as follows: Assuming that the same number of magnetic flux lines pass through both coils, the volts per turn in each coil would be the same. Therefore, the ratio of the secondary voltage to the primary voltage is equal to the ratio of the number of turns in the secondary coil to the number of turns in the primary coil. Equation 1 shows this transformer law.

$$\frac{V_s}{V_p} = \frac{N_s}{N_p} \qquad (1)$$

where

V_s = secondary voltage.
V_p = primary voltage.
N_s = number of turns in secondary coil.
N_p = number of turns in primary coil.

Thus, a step-down transformer has fewer turns in the secondary coil than in the primary coil, and a step-up transformer has more. *Typical transformer problems are solved in Examples A and B by* use of Equation 1.

EXAMPLE A: STEP - UP TRANSFORMER VOLTAGE

Given: The transformer in Figure 1 has 4 turns in the primary winding and 8 turns in the secondary winding. The input voltage to the primary coil (V_p) is 3 volts.

Find: Theoretical output voltage from the secondary coil (V_s).

Solution:

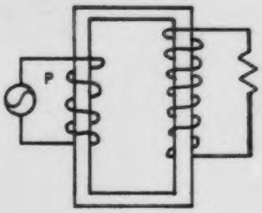
$$\frac{V_s}{V_p} = \frac{N_s}{N_p}$$

$$V_s = \left(\frac{N_s}{N_p}\right) V_p$$

$$= \left(\frac{8 \text{ turns}}{4 \text{ turns}}\right) 3 \text{ V}$$

$$= (2)(3\text{V})$$

$$V_s = 6\text{V}$$



The transformer "steps up" the voltage from 3 volts to 6 volts.

EXAMPLE B: STEP - UP TRANSFORMER WINDING

Given: An electric generator delivers energy at a voltage of 15,000 volts.

Find: The additional number of turns needed in the secondary winding than in the primary winding to step this voltage up to 300,000 V. (Assume 100% efficiency.)

Solution:

$$\frac{V_s}{V_p} = \frac{N_s}{N_p} = \frac{300,000}{15,000} = 20$$

$$N_s = (20)(N_p)$$

The secondary winding will need 20 times as many turns as the primary winding.

As long as the voltage rating of the winding insulation is not exceeded, either winding of a transformer can be used as the primary winding. A step-up transformer can be used as a step-down transformer if its leads are reversed. For example, if the transformer in Example A is reversed by putting 3 V into the side with eight turns, the voltage out of the four-turn winding side will be 1.5 V.

POWER

At first glance, it would seem possible to raise electrical energy to enormous levels through sequential use of step-up transformers. This would be a violation of the conservation-of-energy principle. A transformer does not create energy, it merely transforms it. The rate of transfer is called "power," defined by Equation 2.

$$P = IV \tag{2}$$

where

P = power.
I = current.
V = voltage.

If a transformer increases voltage, it decreases current proportionately, and vice versa; therefore, power is conserved and can be written in a number of useful ways:

Power in = power out

$$I_p V_p = I_s V_s$$

$$\frac{V_s}{V_p} = \frac{I_p}{I_s}$$

If this last equation is considered in light of Equation 1, the transformer law can be written in terms of current, as shown in Equation 3.

$$\frac{I_s}{I_p} = \frac{N_p}{N_s} \qquad (3)$$

where

I_s = current through secondary coil.
I_p = current through primary coil.

The ratio of turns in Equation 3 is inverse to the ratio in Equation 1. The use of Equation 3 for solving problems is shown in Example C.

EXAMPLE C: STEP - UP TRANSFORMER CURRENT

Given: A transformer has 500 primary turns (N_p) and 3000 secondary turns (N_s).

Find: The secondary current (I_s) if the primary current (I_p) is 3 A.

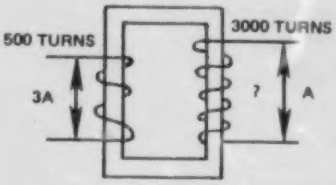
Solution:

$$\frac{I_s}{I_p} = \frac{N_p}{N_s}$$

$$I_s = \left(\frac{N_p}{N_s}\right)$$

$$I_p = \left(\frac{500}{3000}\right) 3\ \text{A}$$

$$I_s = 0.5\ \text{A}$$



3 A is reduced to 0.5 A in a step-up transformer.

To summarize: if the voltage is stepped up, the current is stepped down; if the voltage is stepped down, the current is stepped up.

USE OF OHM'S LAW

Ohm's law, however, states that an increase in voltage corresponds to an increase in current. Is this a contradiction? No, Ohm's law still applies, as the voltages are fixed values. Primary voltage is equal to the voltage of the a.c. source, and the primary current is determined by what happens in the secondary circuit. Secondary voltage is determined by the ratio of the loops in the primary and secondary coils. The ratio is a fixed value; and the secondary current obeys Ohm's law, equal to the induced voltage divided by the load (resistance). The product of the induced voltage and the secondary current is equal to the power in the secondary coil. Since in the ideal case, power in = power out, the primary current then can be determined. Example D explains this further.

EXAMPLE D: OHM'S LAW IN A TRANSFORMER

Given: A transformer has twice as many secondary loops (N_s) as primary loops (N_p). The primary coil is powered by a 100 V (V_p) source and the secondary coil is connected to a 50 Ω load.

Find: (a) Induced voltage (V_s) in the secondary. (b) Current (I_s) in the secondary. (c) Power (P_s) in the secondary. (d) Current (I_p) in the primary. (e) Is this a step-up or step-down transformer.

Solution: (from Equation 1)

(a)

$$\frac{V_s}{V_p} = \frac{N_s}{N_p}$$

$$V_s = \left(\frac{N_s}{N_p}\right) V_p$$

$$V_s = (2)(100\ \text{V}) = 200\ \text{V}$$

(b) $V_s = IR$ (Ohm's law)

$$I_s = \frac{V_s}{R} = \frac{200\ \text{V}}{50\ \Omega} = 4\ \text{A}$$

(c) $P_s = I_s V_s$

$$P_s = (4\ \text{[illegible]}\ 200\ \text{V}) = 800\ \text{Watts}$$

(d) Power in = power out (ideal transformer)

$$I_p V_p = I_s V_s$$

$$I_p = \left(\frac{V_s}{V_p}\right) I_s$$

$$I_p = \left(\frac{200\ V}{100\ V}\right) 4\ A$$

$$I_p = 8\ A$$

or

$$P_p = I_p V_p$$

$$I_p = \frac{P_p}{V_p}$$

$$= \frac{800\ W}{100\ V}$$

$$I_p = 8\ A$$

(e) Since $V_s > V_p$, this is a step-up transformer.

Note: Voltage is stepped up, current is stepped down, power is conserved, and Ohm's law is obeyed.

There are situations in which higher voltages are advantageous, and other situations in which low voltage is more useful. High voltage is necessary in power-transmission lines to keep heat losses to a minimum, and low voltage is necessary in most household appliances to keep them safe. Additionally, low voltage is necessary to yield the high currents needed for welding. These voltages are produced by the proper use of step-up and step-down transformers.

POWER TRANSMISSION

Consider the case of a power plant contracted to deliver 12,000 watts of electrical power (P) over a transmission line that has a resistance of 0.2 ohms (R). Consider further that the power plant unwisely decides to deliver the power at a potential of 120 volts (V). How effective will this power transmission be?

Remembering that P = IV, the current (I) in the transmission line is equal to the following:

$$I = \frac{P}{V} = \frac{12{,}000\ W}{120\ V} = \text{[illegible]}\ A$$

From Ohm's law, $V = IR$, the voltage drop in the transmission line is —

$$V = IR = (100\ A)(0.2\ \Omega) = 20\ V$$

This calculation means that the customer gets the power, not at 120 V, but at 120 V - 20 V or 100 V.

Power loss due to heat in the transmission line varies as the square of the current (I). At 100 A, the power loss due to heat is given thusly:

$$P = I^2R = (100\ A)^2\ (0.2\ \Omega) = 2000\ W$$

The customer only gets 12,000 W - 2,000 W, or 10,000 W, instead of the contracted amount.

Because a transformer was not utilized to step-up the transmission voltage, the power dropped from 12,000 W to 10,000 W, and the voltage dropped from 120 V to 100 V. Both are substantial losses. Example E shows how much more efficient the power transmission can be made by use of a step-up transformer and why transformers are extremely valuable in power transmission.

EXAMPLE E: POWER TRANSMISSION AT HIGH VOLTAGE

Given: A power plant contracted to deliver 12,000 watts of electrical power (P) over a transmission line that has a resistance of 0.2 ohms (R). Through the proper use of a step-up transformer, the potential can be raised from 120 volts to 12,000 volts (V).

Find: (a) Current (I) in the transmission line: (b) Voltage drop along the line. (c) Power loss due to heat. (d) Power and voltage the customer ultimately receives. (Assume that the voltage is stepped back down proportionately at the customer's place of business.)

Solution: (a) $P = IV$

$$I = \frac{P}{V} = \frac{12{,}000\ W}{12{,}000\ V}$$

$$I = 1\ A$$

(b) $V = IR$

$= (1\text{ A})(0.2\ \Omega)$

$V = 0.2\text{ V}$

(c) $P = I^2R$

$= (1\text{ A})^2\ (0.2\ \Omega)$

$P = 0.2\text{ W}$

(d) Power to customer: 12,000 W - 0.2 W = 11,999.8 W

Voltage to customer:
1) Transmitted — 12,000 V - 0.2 V = 11,999.8 V
2) Stepped down — 11,999.8 V ÷ 100 = 119.998 V

The power and voltage drops an insignificant amount.

Ohm's law can be used one step further to develop a resistance equation for transformers. From Equations 1 and 3, expressions for primary voltage and current are derived as follows:

$$V_p = \left(\frac{N_p}{N_s}\right) V_s$$

and

$$I_p = \left(\frac{N_s}{N_p}\right) I_s$$

Since Ohm's law yields $R = V/I$ —

$$R_p = \frac{V_p}{I_p} = \frac{(N_p/N_s)\ V_s}{(N_s/N_p)\ I_s} = \left(\frac{N_p^{\ 2}}{N_s^{\ 2}}\right) R_s$$

Equation 4 shows the relationship between resistance and turns of both coils.

$$\frac{R_p}{R_s} = \left(\frac{N_p}{N_s}\right)^2 \qquad (4)$$

Example F demonstrates the use of this equation.

EXAMPLE F: RESISTANCE OF A TRANSFORMER

Given: A transformer has a secondary resistance (R_S) of 25 ohms, 400 turns in the primary winding, and 100 turns in the secondary winding.

Find: Resistance of the primary winding.

Solution:

$$\frac{R_p}{R_s} = \left(\frac{N_p}{N_s}\right)^2$$

$$R_p = \left(\frac{N_p}{N_s}\right)^2 R_s$$

$$= \left(\frac{400}{100}\right)^2 (25\ \Omega)$$

$$= (16)(25\ \Omega)$$

$$R_p = 400$$

Note: The load resistance of 25 Ω does not include the d.c. resistance of the winding wire which is usually low compared to the load resistance. The load resistance mentioned here is usually called "impedance."

POWER LOSSES

In reality, not all the power put into a transformer is available at the secondary terminals. Some of the power is lost as heat in the transformer windings (copper losses), and some power is lost in the transformer core itself (core losses). Consequently, there are three major sources of power loss in a transformer.

I^{2}R LOSS

Coils of wire have resistance. When current flows through a resistance, power is lost due to heat. This power loss (P), as shown in Example E (c), can be calculated by the equation $P = I^2R$. Thus, power loss due to the winding resistance is called the "I^2R loss."

EDDY-CURRENT LOSS

Current in the primary winding induces current in the core as well as in the secondary winding. This core current, called "eddy current," dissipates power as it flows through the core material. Eddy-current loss can be reduced when the sheets of the core are laminated. If the core is made of thin sheets of magnetic material separated by insulating varnish, the length of the eddy current paths is reduced, thereby reducing the power loss. In some transformers, to reduce eddy current, the core is made of magnetic powder suspended in a ceramic-like, insulating material.

HYSTERESIS LOSS

Energy is required to reverse the direction of core magnetization every half cycle of the input frequency. This energy, called "hysteresis," is lost to heat like I^2R and eddy-current losses, and is controlled by the use of iron cores having low hysteresis susceptibility. Hysteresis losses are greater at higher frequencies, since the core magnetization must reverse more frequently. In an effort to reduce hysteresis loss at very high frequencies, an iron core is not used. The secondary and primary coils are simply wound on a nonmagnetic, nonconducting core.

Losses like these in a well-designed transformer do not normally exceed two or three percent; however, at these losses the power out of a transformer will always be less than the power in. The ratio of power out to power in is called the "efficiency." In all of the discussion and example problems to this point, an ideal efficiency of 100 percent has been assumed (power out = power in). For the sake of accuracy, the true transformer efficiency should be included.

In Figure 1, a transformer having windings on opposite sides of the transformer core was shown. In practice, the core is shaped with a middle core piece, as shown in Figure 3, with the secondary coil wound on top of the primary.

Transformers can be wound with many secondary windings or with several connection taps on one secondary winding to suit the application.

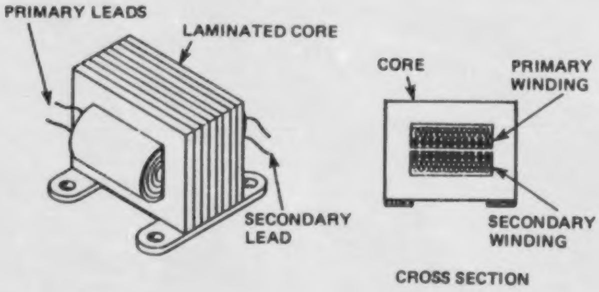


FIG. 3

SUMMARY

The electrical transformer steps up or steps down voltage through the process of magnetic induction. The a.c. transformer consists of (1) a coil connected to an a.c. source, (2) another (secondary) coil attached to a load, and (3) a soft iron core. The number of wire turns in each coil determines the amount of voltage change; the coil with the larger number of turns is capable of stepping up the voltage, and vice versa.

LABORATORY

MATERIALS

Laminated iron transformer core.
Spool of conducting wire.
Voltmeter or volt-ohm-milliammeter (VOM) (Triplett 630 or Simpson 260).
Low voltage (0-12 V) a.c. power supply.
1 kΩ resistor.
Assorted test leads.

PROCEDURE

1. Form a primary coil for a transformer by winding 24 turns (N_p) of conducting wire around one part of a laminated iron core.

2. Form a secondary coil around the same core, but on opposite side, by winding 6 turns (N_s) of conducting wire.

3. Connect transformer to a circuit, as shown in Figure 4.

4. Plug in a.c. power supply, and adjust it to deliver 6 volts (V_p) to primary coil. Record the voltmeter reading (V_s) in the Data Table. Compare this experimental value of V_s with the value calculated using the data given for V_p, N_p and N_s. Repeat with V_p = 10 volts.

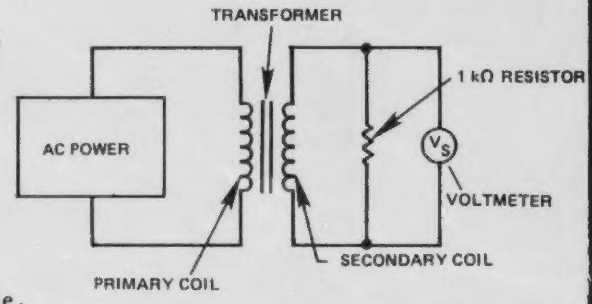


FIG. 4

5. Unplug the power supply. Rewind the secondary coil of the transformer with 12 turns of conducting wire. Repeat steps 3 and 4.

6. Increase the number of turns in the secondary coil to 36 and 48 turns. Record all data in the Data Table.

7. Using calculated values of secondary voltage (V_s), calculate secondary current (I_s), secondary power (P_s), and primary current (I_p). Assume the transformer to be 100% efficient (power in = power out).

8. Why is an a.c. current (or at least a pulsating current) necessary for transformer operation?

DATA TABLE

Trial No.	Primary Turns (N_p)	Secondary Turns (N_s)	Primary Voltage (V_p) Volts	Secondary Voltage (V_s) (Calc) Volts	Secondary Voltage (V_s) (Meas) Volts	Secondary Current (I_s) Amperes	Secondary Power (P_s) Watts	Primary Current (I_p) Amperes
1	24	6	6					
			10					
2	24	12	6					
			10					
3	24	36	6					
			10					
4	24	48	6					
			10					

SELF CHECK

1. A transformer has $N_p = 400$ turns, $N_s = 900$ turns, $V_p = 8$ V and $I_p = 0.4$ A. Calculate I_s and V_s.

2. A transformer has $N_p = 1000$ turns, $N_s = 200$ turns, $I_p = 0.8$ A, $V_p = 10$ V. Calculate I_s and V_s.

3. A transformer has $V_p = 12$ V and $V_s = 3$ V. Calculate the turns ratio.

4. A transformer has $I_p = 2$ A and $I_s = 0.1$ A. Calculate the turns ratio.

5. Briefly describe the three major sources of transformer power loss.

6. A transformer has N_p = 3600 turns, N_s = 600 turns, and R_s = 50 Ω. Calculate R_s.

7. If a transformer has N_p = 4000 turns, N_s = 2000 turns, and R_s = 50 Ω, what are the primary and secondary currents (I_p and I_s) if V = 100V? Assume 100% efficiency.

ENERGY CONVERTORS

APPLICATION MODULES

		Page
9F1	Turbines	60
9F2	Fluid Pumps	74
9F4	Fans and Blowers	89
9F5	Vacuum Pumps	101
9E1	Electric Generators	117
9E2	Alternators	129
9E3	Electric Motors	143
9E4	Solenoids and Relays	158
9E5*	Photovoltaic Materials	168
9T2	Thermoelectric Generators	180
9T3	Solar Collectors	197

* Identifies concept support modules that introduce new physics.

9F1

TURBINES

The turbine, a prime mover in technology today, is a device that converts fluid energy into mechanical rotational energy. Gas turbine-powered automobiles and trucks are a reality, and gas-turbine turboprop engines are in daily use in aviation. Modern ocean liners are propelled by marine turbines. Various steam turbine-equipped generating plants (including nuclear plants) produce approximately 80% of the electricity in the United States. Water-powered turbines produce nearly 20% of America's electrical power.

In this module, the basic characteristics and operating principles of a turbine are presented. Energy transformations will be traced diagrammatically for gas, steam, wind and water turbines. In the laboratory the student will determine the efficiency of a small model water turbine.

OBJECTIVES

Upon completion of this module, the student should be able to:

- List two advantages of a turbine, compared to a reciprocating engine.
- Name and give the function of two major elements of a turbine.
- In one or two sentences, define the following terms:

  Impulse turbine.
  Reaction turbine.
  Staging.

- Calculate the input fluid power and efficiency of a turbine, given the input pressure, the volume flow rate and the output power.

- Given the appropriate equipment, calculate the input power, the output power and the efficiency of a model water turbine.

DISCUSSION

Energy conversions occur through a device called an energy convertor. A source of energy is provided to the convertor in initial form. As the energy passes through the convertor, it is converted to a desired output form of energy, as shown in Figure 1. Unfortunately, some of the input energy is also converted to a waste form of energy such as heat, which represents a loss of usable energy in the conversion process.

A well-known energy convertor is the turbine, which consists of two major elements, a high-energy working fluid and a rotor. The input energy to the turbine is provided by the high-energy fluid (combustion gases, steam, wind or a flowing fluid such as water). The flowing medium has an initial high energy content as it enters the turbine. It flows through the turbine and causes the rotor to spin. The rotor, in effect, extracts energy from the moving fluid and converts the extracted fluid energy to a mechanical rotational form. The mechanical rotational energy is transmitted along the rotor shaft where it is available as useful output energy. These elements of a turbine operation and energy flow are illustrated in Figure 2.

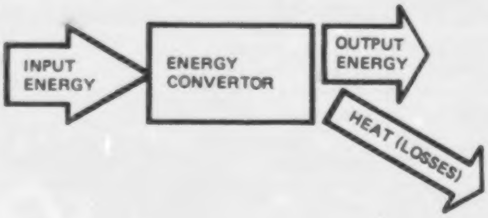

INPUT ENERGY
ENERGY CONVERTOR
OUTPUT ENERGY
HEAT (LOSSES)

FIG. 1

The efficiency of an energy transformer is given by Equation 1.

$$\eta = \frac{E_{out}}{E_{in}} (100) \qquad (1)$$

where
η = % efficiency.
E_{in} = energy input.
E_{out} = energy output.

Since output energy and input energy are developed and used over the same time interval, an alternate form of Equation 1 is given by Equation 2.

$$\eta = \frac{P_{out}}{P_{in}} (100) \qquad (2)$$

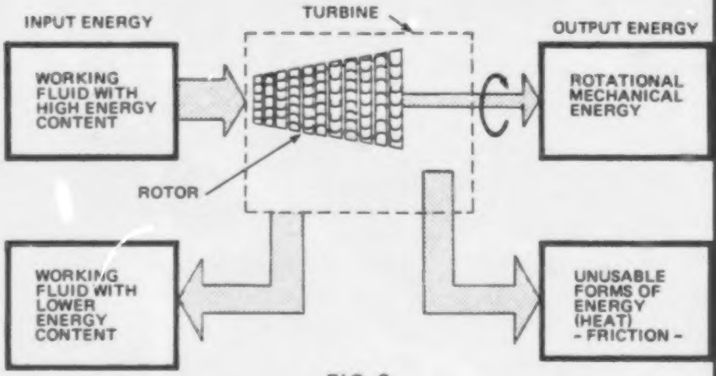

INPUT ENERGY
TURBINE
OUTPUT ENERGY
WORKING FLUID WITH HIGH ENERGY CONTENT
ROTATIONAL MECHANICAL ENERGY
ROTOR
WORKING FLUID WITH LOWER ENERGY CONTENT
UNUSABLE FORMS OF ENERGY (HEAT) - FRICTION -

FIG. 2

where

P_{in} = power input.

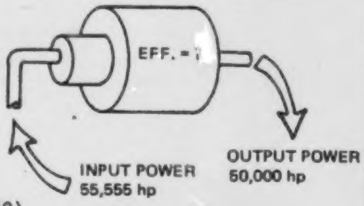
P_{out} = power output.

Equation 2 is used to find efficiency in Example A. High efficiency is one of the distinct advantages of a turbine over reciprocating engines since the latter wastes much energy starting and stopping heavy pistons. (See Example A.) Turbines are designed to rotate smoothly at high rpm with little vibration and no change in direction of the rotor. Since there is little vibration, turbines also may be lighter in weight than other engines.

EXAMPLE A: EFFICIENCY CALCULATIONS

Given: A marine turbine which develops 55,000 horsepower at the rotor shaft and requires an input power of 55,555 horsepower.

Find: The % efficiency.



Solution:

$$\eta = \frac{P_{out}}{P_{in}} \ (100)$$

$$= \frac{(50{,}000 \ \cancel{hp})}{(55{,}555 \ \cancel{hp})} \ (100)$$

$$\eta = 90\%.$$

The modern water turbines used in generating electricity convert the available potential heat energy of water into electrical energy at efficiencies as high as 94%. Some of these units develop as much as 700,000 horsepower. The efficiency of steam turbines can be improved if the rotor is constructed with series of axial stages having alternate rows of rotating and fixed blades, as shown in Figure 3. Each stage extracts energy from the fluid flowing into it from the previous stage. Turbines of this type are known as impulse-reaction turbines. The term "impulse" refers to the process by which the available heat energy in steam is transformed into kinetic energy as the steam jet flows from the nozzles. This kinetic energy is transformed to work done on the rotor by impinging on a rotating row of blades. In pure impulse turbines, such as some water turbines, the working fluid undergoes little pressure change passing through a stage. The term "reaction" turbine is applied to the type of turbine which derives an additional thrust by directing the working fluid through a curved blade, thereby changing the fluid's direction. An additional thrust is gained from the change in momentum of the working fluid. Most actual turbines are a combination of impulse and reaction principles. Practical turbine systems for converting fluid energy to mechanical energy in electrical generating plants often consist of two or three turbines mounted on the same shaft. Thus, the exhaust steam from a high-pressure turbine may be used to power another turbine designed for lower pressure operation.

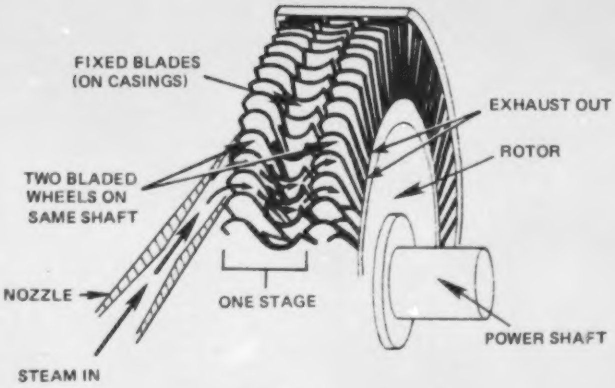

FIXED BLADES (ON CASINGS)
EXHAUST OUT
ROTOR
TWO BLADED WHEELS ON SAME SHAFT
NOZZLE
ONE STAGE
POWER SHAFT
STEAM IN

FIG. 3

MODEL WATER TURBINE

IMPULSE-REACTION TURBINE

Figure 4 shows the variables that must be measured in order to calculate the input power to a water turbine. The input power to the water turbine is given by Equation 3.

$$P_{in} = p\frac{V}{t} \qquad (3)$$

where

P_{in} = power input.
p = pressure.
V = volume.
t = time.

The use of Equation 3 is illustrated in Example 8.

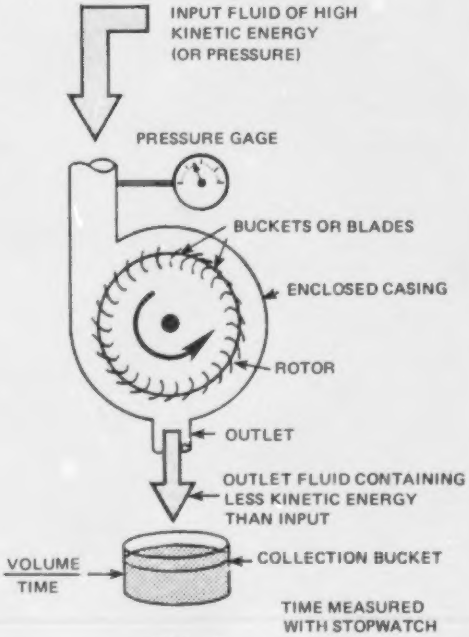

INPUT FLUID OF HIGH KINETIC ENERGY (OR PRESSURE)
PRESSURE GAGE
BUCKETS OR BLADES
ENCLOSED CASING
ROTOR
OUTLET
OUTLET FLUID CONTAINING LESS KINETIC ENERGY THAN INPUT
VOLUME / TIME
COLLECTION BUCKET
TIME MEASURED WITH STOPWATCH

FIG. 4

EXAMPLE B: CALCULATION OF INPUT POWER

Given: A water turbine in which the pressure of the incoming fluid is 50 lb/in^2 and the volume of water collected is 0.25 cubic feet (roughly 2 gal.) in a 2-minute period of time.

Find:

(a) Calculate the input power, and express the answer in horsepower.

(b) Convert the given data to metric units, and calculate the input power in watts.

Solution:

$$p = \text{pressure} = \left(50\ \frac{\cancel{\text{lb}}}{\cancel{\text{in}^2}}\right)\left(\frac{6895\ \text{N/m}^2}{1\ \cancel{\text{lb/in}^2}}\right) = 344{,}750\ \frac{\text{N}}{\text{m}^2}$$

$$V = \text{volume} = (0.25\ \cancel{\text{ft}^3})\left(\frac{1\ \text{m}^3}{35.315\ \cancel{\text{ft}^3}}\right) = 0.0071\ \text{m}^3$$

$$t = \text{time} = (2\ \cancel{\text{min}})\left(\frac{60\ \text{s}}{1\ \cancel{\text{min}}}\right) = 120\ \text{s}$$

(a) English System:

$$P_{in} = p\,\frac{V}{t}$$

$$= \left(50\ \frac{\text{lb}}{\cancel{\text{in}^2}}\right)\left(\frac{144\ \cancel{\text{in}^2}}{1\ \text{ft}^2}\right)\left(\frac{(0.25\ \text{ft}^3)}{120\ \text{s}}\right) = \frac{15\ \text{ft}\cdot\text{lb}}{\text{s}}$$

$$= \frac{15\ \frac{\text{ft}\cdot\text{lb}}{\text{s}}}{\left(\frac{550\ \frac{\text{ft}\cdot\text{lb}}{\text{s}}}{1\ \text{hp}}\right)} \qquad \left(\text{Recall:}\quad 550\ \frac{\text{ft}\cdot\text{lb}}{\text{s}} = 1\ \text{hp}\right)$$

$$P_{in} = 0.027\ \text{hp}$$

(b) Metric System:

$$P_{in} = p\,\frac{V}{t}$$

$$= \left(344{,}750\ \frac{\text{N}}{\cancel{\text{m}^2}}\right)\left(\frac{(0.0071\ \cancel{\text{m}^3})}{(120\ \text{s})}\right)$$

$$P_{in} = 20.4\ \text{watts}$$

Note: Metric units are more appropriate for power measurements of small laboratory water turbines. A useful conversion is 1 hp = 746 watts.

OUTPUT POWER OF A LABORATORY WATER TURBINE

The output power from the laboratory water turbine is available at the rotating shaft in the form of mechanical rotational, kinetic energy per unit time. The amount of output power will be less than the input power because of losses in the energy transformation process. In order to measure mechanical rotational power, another energy transformation is used. A device called a prony brake is constructred in such a manner that when applied to the rotating shaft, a frictional force (a braking load) is applied to the shaft; and the available power is converted to dissipated power in the form of heat developed per unit time and any remaining mechanical rotational power. This power transformation is illustrated by Equation 4.

$$\overbrace{\underset{\underset{\text{No load condition.}}{\uparrow}}{\begin{bmatrix}\text{Available}\\ \text{mechanical}\\ \text{rotational}\\ \text{power}\end{bmatrix}}}^{\text{FROM THE TURBINE}} = \overbrace{\begin{bmatrix}\text{braking load}\\ \text{power, less}\\ \text{loss in form}\\ \text{of heat}\end{bmatrix} + \underset{\underset{\text{Evidenced by remaining rpm at the shaft.}}{\uparrow}}{\begin{bmatrix}\text{remaining}\\ \text{mechanical}\\ \text{rotational}\\ \text{power at}\\ \text{the shaft}\end{bmatrix}}}^{\text{AT THE PRONY BRAKE}} + \begin{matrix}\text{other}\\ \text{losses}\end{matrix} \qquad (4)$$

When the prony brake is adjusted in order that the turbine almost stalls (the output shaft ceases to rotate), then all the available mechanical rotational, kinetic power from the water turbine would be converted to braking load power (plus other, hopefully negligible, losses).

The prony brake consists of a friction belt that wraps around a smooth drum attached to the output shaft of the water turbine. The ends of the belt are attached to spring scales, as shown in Figure 5. The braking load (frictional force) applied at the surface of the drum is the difference in the spring scale readings $(F_1 - F_2)$. A mechanical revolution counter attached to the rotating shaft enables revolutions per unit time to be determined.

Braking load power is given by Equation 5.

$$P_{\text{load}} = (F_1 - F_2)\,\frac{\pi d N}{t} \qquad (5)$$

where

P_{load} = breaking load power.
F_1 = highest spring scale reading.
F_2 = lowest spring scale reading.

π = 3.14.
d = drum diameter.
N = number of revolutions.
t = time in seconds for N revolutions.

Equation 5 is used to solve a problem in Example C.

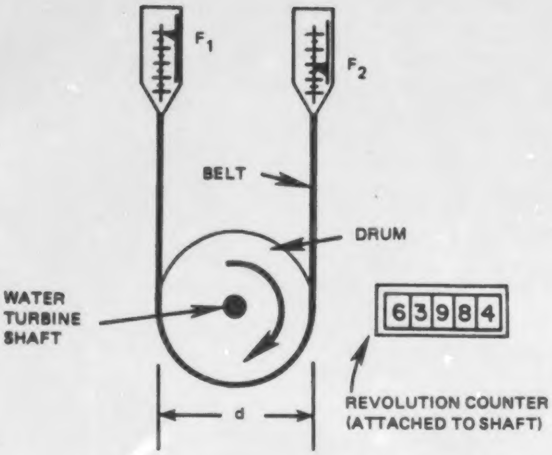


FIG. 5

EXAMPLE C: CALCULATION OF BRAKING LOAD POWER

Given: A prony brake turns 180 revolutions in 3 minutes when the spring scale readings are 15.68 N and 0.98 N respectively and the diameter of the friction drum is 8 cm.

Find: The braking load power (in watts) developed.

Solution: From Equation 5:

$$P_{load} = (F_1 - F_2)\,\frac{\pi d N}{t}$$

$$= (15.68 - 0.98)(3.14)(0.08\ \text{m})\left(\frac{180\ \text{rev}}{180\ \text{s}}\right)$$

$$P_{load} = 3.69\ \text{watts}$$

Efficiency of a laboratory turbine is calculated in Example D. It is low compared to the 94% efficiency possible with some large turbines. The laboratory turbine is a single-stage unit that extracts only a small percentage of the power flowing through it.

EXAMPLE D: CALCULATION OF EFFICIENCY

Given: A laboratory water turbine with power input of 20.4 watts and braking power output of 3.69 watts.

Find: % efficiency.

Solution:

$$\eta = \frac{P_{out}}{P_{in}} (100)$$

$$= \frac{(3.69 \text{ watts})}{(20.4 \text{ watts})} (100)$$

$$\eta = 18\%$$

SUMMARY

The turbine converts fluid energy into mechanical rotational energy. It is one of the more efficienct devices available, and is used as a power source for trucks, ships, aircraft, electrical power plants, and other equipment. The major sources of input energy to the turbine are gas, steam, wind or water.

EFFICIENCY OF A WATER TURBINE

MATERIALS

Water turbine, equipped with pressure gage, 0 to 60 psi.
Prony brake with scales and mechanical revolution counter.
    (Available from Central Scientific Company or equal supplier.)
A two-gallon catch bucket.
Stopwatch.
Sink.

PROCEDURE

The model water turbine used for this laboratory experiment develops such small power that calculations should be done in the metric system of units and measurements. The reason for this suggestion are indicated by the results of Example B. Data Table 1 should be completed using both the English and the metric systems.

1. Arrange the water turbine and its associated equipment in the manner shown in Figure 6. Note: Some of this procedure requires two persons.

2. Check system for leaks, since a 30 lb/in^2 to 60 lb/in^2 pressure is required.

3. Arrange catch bucket in such a manner as to allow it to catch all water that passes through turbine in a time measured with stopwatch.

4. Shut off water, and dry catch bucket. Move catch bucket out of way of exit tube.

5. Measure diameter of output pulley drum, and record this data in Data Table 1.

6. Adjust position of prony brake scales to produce a small tension in brake band. (The scale readings will normally be the same with turbine stopped.)

7. Turn on water supply to turbine, and record water pressure in Data Table 1, Trial 1.

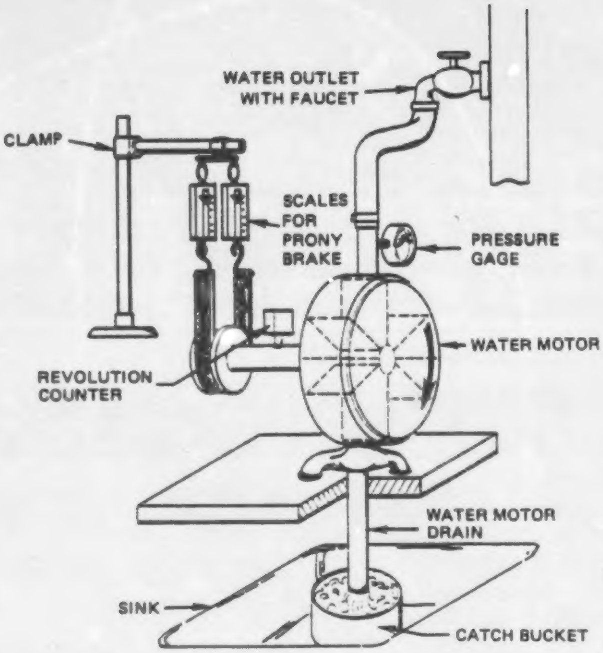


FIG. 6

8. Simultaneously start the stopwatch, slide catch bucket underneath exit tube, and engage mechanical revolution counter at output side of turbine. In a measured time, catch as much water in catch bucket as can be handled reasonably without spilling.

9. During trial, record prony brake scale readings, F_1 and F_2 in Data Table 1.

10. Simultaneously, stop stopwatch, remove catch bucket from underneath exit tube and disengage the mechanical revolution counter. Record in Data Table 1 number of revolutions indicated by counter during trial and time interval shown on stopwatch.

11. Measure volume of water collected during trial, and record in Data Table 1.

12. Repeat steps 7 through 11 for increasing tensions in prony brake band. Perform at least 6 more trials with increasing tensions in brake band, ending with a trial in which turbine is almost stalled by braking load placed on it.

13. Convert the collected data in Data Table 1 to metric units and record the values in the indicated columns of Data Table 1. (Table 1 gives metric conversion factors.)

14. Using data taken from Data Table 1, calculate input power to turbine in watts for trials 1 through 7 of the experiment; record in Data Table 2.

15. Using data taken from Data Table 1, calculate output power from turbine in watts for trials of the experiment, and record in Data Table 2.

16. Calculate and record % efficiency for each trial, using Equation 2:

$$\eta = \frac{P_{out}}{P_{in}} (100)$$

17. Calculate revolutions per second and braking load for each trial. Record in Data Table 2.

18. Why would a multi-stage turbine be more efficient than this single-stage turbine, even when the inlet water pressure and flow rate are the same?

DATA TABLE 1

WATER TURBINE

TRIAL	WATER PRESSURE (GAGE) lb/in²	n/m²	F_1 lb N	F_2 lb N	REVOLUTIONS N	TIME INTERVAL t(s)	VOLUME OF WATER (m³/s)
1							
2							
3							
4							
5							
6							
7							

Diameter of pulley drum: ______ meters.

TABLE 1

CONVERSION FACTORS FOR METRIC UNITS

1 lb = 4.448 N,	1 m = 39.37 in
$1\ \frac{\text{lb}}{\text{in}^2}$ = 6.895 KPa,	$1\ \text{m}^3 = 35.315\ \text{ft}^3$
$= 6895\ \frac{\text{N}}{\text{m}^2}$.	

DATA TABLE 2

TRIAL	POWER INPUT (WATTS)	POWER OUTPUT (WATTS)	EFFICIENCY	REVOLUTION PER s	BRAKING LOAD $(F_1 - F_2)$ (NEWTONS)
1					
2					
3					
4					
5					
6					
7					

SELF CHECK

1. Describe two characteristics of turbines that result in advantages over internal combustion engines.
2. Name and give the function of the two major elements of a turbine.
3. Describe the differences between an impulse turbine and a reaction turbine.
4. Describe staging in an impulse-reaction turbine.
5. A turbine has an input pressure of 100 psi, a flow rate of 400 ft^3 per minute and an efficiency of 75%. Calculate the input power and output power of the turbine in horsepower and kilowatts.

6. How many joules of energy are developed in 5 minutes by a water turbine that when connected to a prony brake turns the brake 100 revolutions in 2 minutes when the spring scale readings are 27 N and 2 N respectively and the diameter of the friction drum is 12 centimeters?

7. Find the efficiency of a water turbine that uses 25 J of input energy in 6 seconds and outputs 60 J of energy every minute.

8. What is the pressure of an incoming fluid if the volume collected in a 6 minute time period is 5 cubic feet and the input power is 30 ft•lb/s?

9. A prony brake is to be tested for accuracy by attaching it to a water turbine of known efficiency (27%). If the input power to the turbine is 20 watts, the prony brake drum has a diameter of 9 cm and the brake turns 220 revolutions in 4 minutes, what is the expected difference in highest and lowest spring scale readings?

9F2

FLUID PUMPS

Fluid pumps are energy convertors that transform mechanical energy into fluid energy. Fluid energy supplied by pumps is utilized either to circulate fluid or to pressurize it for use in hydraulic fluid motors or cylinders. Fluid pumps may be classified by a number of characteristics, but the most fundamental is displacement. The non-positive displacement, or hydrodynamic, pump is usually of the centrifugal type and is used to move large volumes of fluid at low pressures. Positive displacement pumps produce high pressure because the fluid in them is propelled by tightly-fitted mechanical components such as gears, vanes, or pistons.

This module discusses the principles, characteristics, and applications of the hydrodynamic pump and several types of positive displacement pumps. In the laboratory, the student will examine the characteristics of a hydrodynamic pump.

OBJECTIVES

Upon completion of this module, the student should be able to:

- State four characteristic differences between hydrodynamic pumps and positive displacement pumps.
- State the most common application of the hydrodynamic pump.
- State the most common application of the positive displacement pump.
- Draw and label diagrams of the following types of pumps; show direction of motion for moving parts and direction of

fluid flow:

Centrifugal.
External gear.
Unbalanced vane.
Axial piston.

- Explain with the use of diagrams how unbalanced vane pumps and radial and axial piston pumps can be used as variable volume pumps.

- Given the appropriate equipment, make the necessary measurements, and draw the pump characteristic curves for a centrifugal pump.

PUMP PRINCIPLES

Fluid pumps are energy convertors that transform mechanical energy from a motor into fluid energy. Mechanical force is exerted at the outlet by a motor, causing a partial vacuum at the inlet, as demonstrated by the piston pump in Figure 1. When the piston (1) is moved by mechanical force to the left, a partial vacuum is created in the chamber (4). The drop in pressure closes the outlet check valve (3) and opens the inlet check valve (2), permitting fluid to fill the chamber. As the piston moves to the right, the pressure it exerts upon the fluid causes the inlet check valve (2) to close and the outlet check valve (3) to open and thereby to allow flow at the outlet.

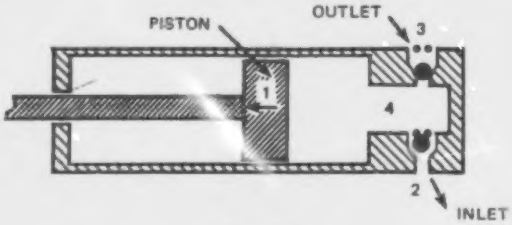


FIG. 1

A potential danger in pumps is "cavitation," which refers to the "cavity" of gases produced when the pressure at the inlet of the pump is so low that fluid vaporizes. Such vapor is carried by the pump through the outlet, where it rapidly compresses under pressure, causing shock waves so powerful that pump components can be damaged. Auxiliary pumps often are used to increase inlet pressure.

THE NON-POSITIVE DISPLACEMENT (HYDRODYNAMIC) PUMP

The centrifugal pump, illustrated in Figure 2, is the most common hydrodynamic type. The principle behind this pump is centrifugal force produced when an object is forced in a circular motion. The object tends to move in a straight line away from the center of the circle. In the centrifugal pump, fluid enters the inlet at the center of the pump into a rotating, finned impeller that spins the fluid so rapidly that it moves, in a straight line, through the outlet. (Centrifugal force is not to be confused with centripetal force, or "force pulling toward the center.")

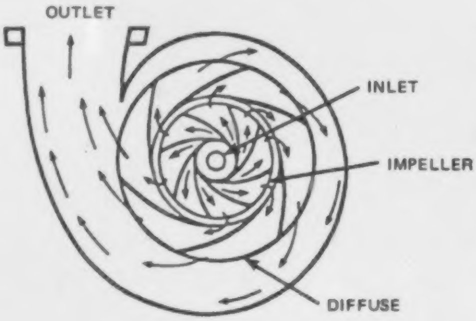


FIG. 2

There is always some "slippage" within the centrifugal pump. Slippage is the fluid not forced through the outlet of the pump although contacted by the pump components. Even at a constant speed the centrifugal pump does not always displace, or move, the same amount of fluid, because the amount of suction head and pressure head determine how much fluid will be moved by the pump. "Suction head" is the amount of pressure the pump must exert upon a fluid of a given density to raise it a given height to the pump. "Pressure head" is the back pressure caused by the equivalent height against which the fluid must be pumped, or the resistance of the load.

When head pressure is great, slippage will be greater and displacement less. Even if the outlet of a centrifugal pump is blocked, slippage prevents damage to pump components. The pump displayed in Figure 2 is an improved design with a diffuser that reduces slippage and increases efficiency. The centrifugal pump produces low pressure but high volumes. It must be primed with fluid before it will start; that is, it must contain enough fluid within its housing to begin pumping action.

A characteristics graph, similar to that given in Figure 3, often is used to determine the size of centrifugal pump required for a given job. Pump size depends upon volume of fluid to be moved, time required to move it, suction and head pressure to be overcome, motor horsepower, and pump efficiency. Three characteristics for a centrifugal pump are plotted against volume flow rate, which is represented in gallons per minute at the bottom of the graph. Notice that each curve is associated with different vertical scales. Delivery pressure, represented by the top curve, drops as the flow rate increases. Overall efficiency rises to a maximum and then drops. Brake horsepower requirement for the pump motor increases as the total head and flow rate increase. Example A illustrates the use of a pump characteristic graph in plotting flow rate and efficiency.

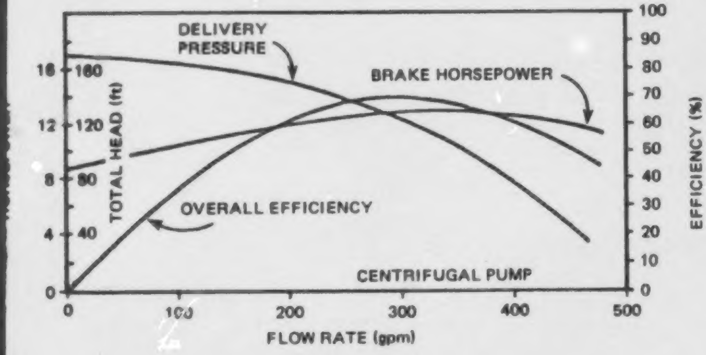


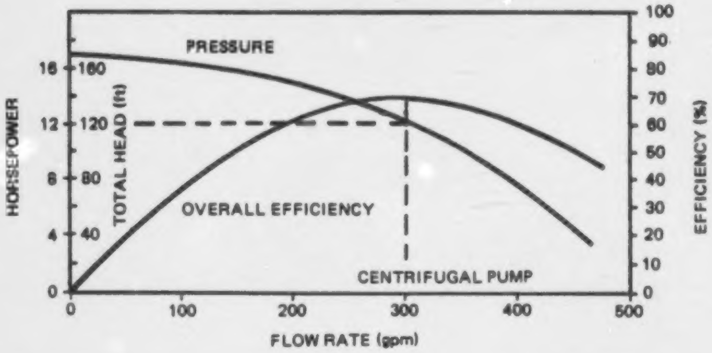
FIG. 3

EXAMPLE A: USE OF PUMP CHARACTERISTICS GRAPH

Given: The below graph is for a particular pump.

Find: The flow rate and % Efficiency of this pump when there is a total head of 120 ft.

Solution: Locate total head in left vertical column. Extend line to right until it intersects Pressure curve, and draw line, at 90°, to horizontal axis (flow rate), which gives a flow rate of approximately 310 gallons per minute. Extend vertical line upward until it intersects the overall efficiency curve. From this intersection, draw horizontal line to right margin, which is the % Efficiency line. In this case the efficiency is 70% (peak).



The most frequent application of the centrifugal pump is the circulation of large volumes of fluid at low pressure. Applications include use in automobile water pumps and lubricating pumps for machinery and shop processes.

POSITIVE DISPLACEMENT PUMPS

The term "positive displacement" means that a specific volume of fluid is moved, or displaced, each time the pump makes a revolution, regardless of pump speed. There is little slippage in this kind of pump, because the fluid is moved by tightly-fitted components, such as gears, vanes, or pistons. Greater pressures, but lower volumes, are achieved than with the centrifugal pump. If the outlet is blocked, pump damage will result, unless safety valves are built into the system. No priming is required because of little slippage. Hydraulics is the most common application of the positive displacement pump.

Positive displacement pumps are classified also by delivery and motion. Delivery of a fluid by a pump may be either constant of variable. The constant-delivery pump displaces the same amount of fluid in a given time at the same speed. Fluid from the constant-delivery pump drives machines not requiring a speed change. The variable-delivery pump contains design features that enable its output volme to be changed manually or automatically without changing the speed of the power source. Variable-delivery pumps are used in such applications as conveyor systems, continuous-processing equipment, and constant-tension devices, in which variable delivery fluid rate is necessary.

Motion is either rotary or reciprocating. Rotary pumps utilize components that move in a circular direction. Reciprocating pumps contain pistons that move backwards and forwards, as was shown earlier in Figure 1.

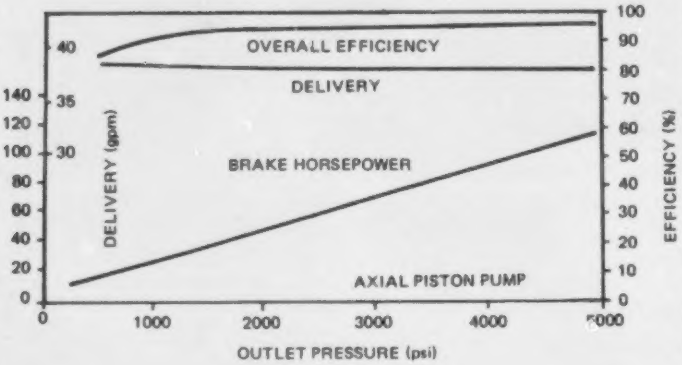


FIG. 4

Figure 4 consists of pump characteristics for use with positive displacement pumps. In this case, pump outlet pressure in psi, is represented at the lower part of the graph. Since there is almost no slippage, the delivery flow rate is constant at all pressures. Overall efficiency rises from about 80% at very low pressure (for example, at 100 psi) to over 95% from 1000 psi to capacity. The only characteristic that changes significantly is that for required brake hp, which increases linearly with pressure.

TYPES OF POSITIVE DISPLACEMENT PUMPS

The three types of positive displacement pumps are (1) gear, (2) vane, and (3) piston.

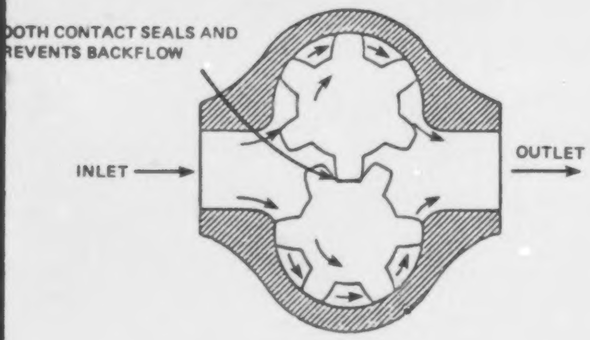


FIG. 5

GEAR PUMPS

Figure 5 illustrates an external gear pump. Both external gears (spur gears) have teeth around their outside surfaces. One is a motor-driven gear, which in turn drives the second gear, causing a tight seal to exist between the meshing teeth to prevent slippage. Both gears are fitted tightly into a housing. As the grooves in the gears collect fluid at the inlet and rotate, the fluid is sealed in the grooves as the spurs contact the housing. The gears rotate to the outlet side of the housing, where they release their fluid.

Internal gear pumps have one large, internal, ring gear with teeth on its inside surface, as in Figure 6. An external gear (with teeth on its external surface) fits inside, its teeth fitting closely into the internal gear teeth on one side (top left side in drawing). Fluid is drawn into the grooves of both moving gears at the inlet as the outer gear ceases contact with the inner in that area. Fluid is held in the grooves by a stationary crescent until both gears rotate to the outlet, where the external teeth again contact the inner, forcing the fluid through the outlet. Internal gear pumps are capable of greater pressures than external ones.

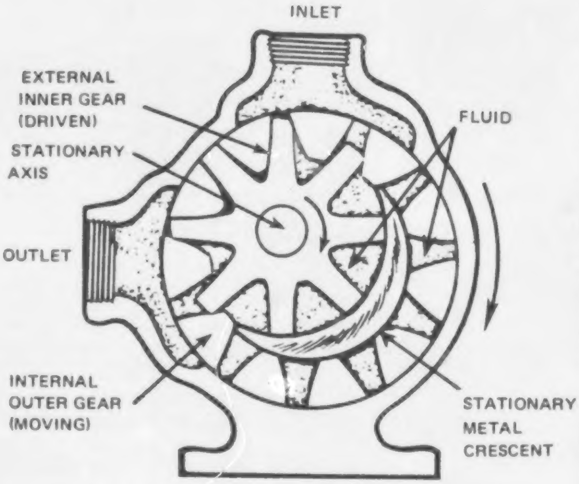

INLET
EXTERNAL INNER GEAR (DRIVEN)
FLUID
STATIONARY AXIS
OUTLET
INTERNAL OUTER GEAR (MOVING)
STATIONARY METAL CRESCENT

FIG. 6

Similar to the external gear pump is the lobe pump (Figure 7). Both lobe elements are driven externally but do not actually make contact, which causes less wear but more slippage. The lobe pump has a large volume output per revolution.

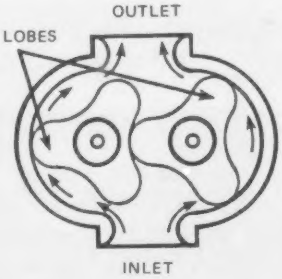


FIG. 7

VANE PUMPS

Vane pumps are either unbalanced (with variable or unvariable volume capacity) or balanced. As indicated in Figure 8, the unbalanced vane pump of the unvariable type consists of a circular rotor containing radial sliding vanes held against the housing by spring pressure behind each vane and by the rotary motion. Because the rotor is unbalanced (off center) in the housing, fluid is trapped at the inlet between the housing and the vanes, as in the grooves of a gear pump. As the rotating vanes deliver fluid to the outlet, they are pushed inward by the housing, causing the fluid to be compressed between the rotor and the housing and to flow through the outlet. The vane pump is used on machine tools because of its reliability and low initial cost.

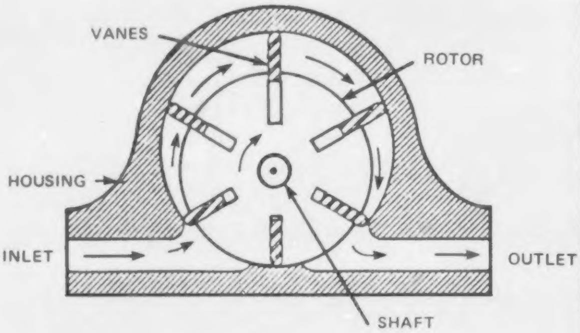


FIG. 8

Figure 9 demonstrates an unbalanced vane pump whose output can be varied externally when the position of either of the two movable cylinder casings is changed. Fluid flow is reversible by movement of the casings.

The balanced vane pump (Figure 10) is not capable of variable-volume output, although its design reduces bearing pressure and wear found in the unbalanced type. The balanced vane pump is designed with two outlets and two inlets. It contains two opposite areas of equal fluid movement, balanced on opposite sides of the housing.

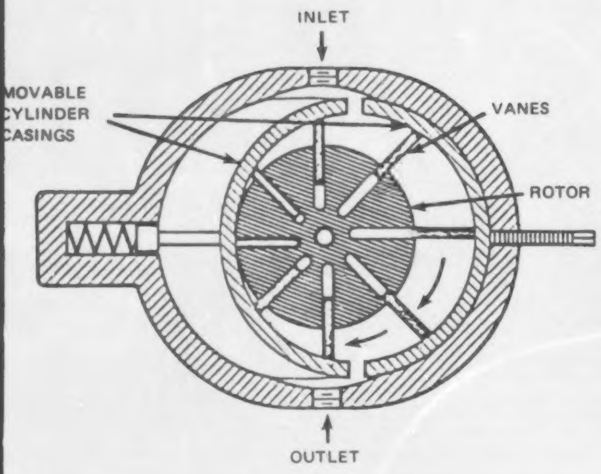

INLET
MOVABLE
CYLINDER
CASINGS
VANES
ROTOR
OUTLET

FIG. 9

PISTON PUMPS

Both radial and axial piston pumps contain reciprocating pistons. Figure 11 displays a radial piston pump, termed "radial" because piston motion is along a radius. The rotor is off center, or unbalanced, as in the unbalanced vane pump. However, instead of fluid being carried in a circular motion between vanes, it is forced into circular pistons as they move outward from the rotor against the housing casing. Springs and rotary motion cause the pistons to press against the casing. When the rotor containing the pistons rotates against the unbalanced casing, the pistons are forced inward and

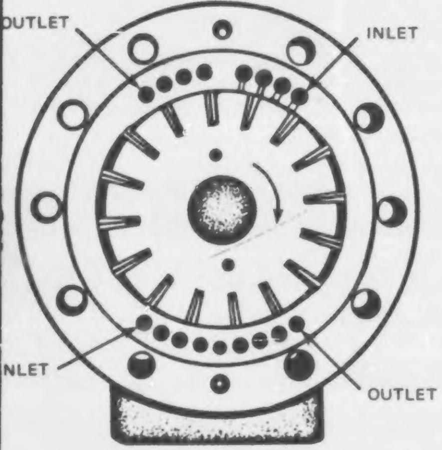

OUTLET
INLET
NLET
OUTLET

FIG. 10

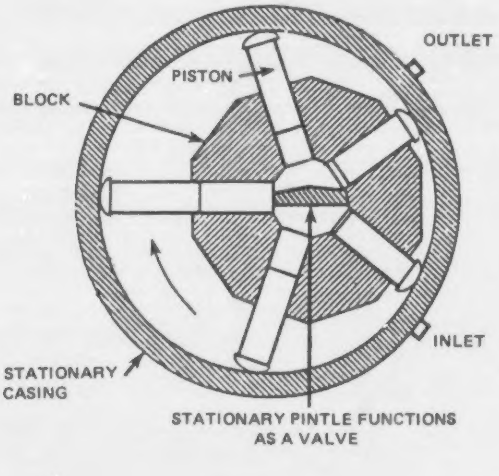

OUTLET
PISTON
BLOCK
INLET
STATIONARY
CASING
STATIONARY PINTLE FUNCTIONS
AS A VALVE

FIG. 11

outward; and the fluid in them passes through a stationary pintle that is open just at the instant when the fluid is forced through at the center of the rotor. Most radial pumps have the variable-volume capacitymade possible by movable casings.

Axial piston pumps operate upon the principle shown in Figure 12. These pumps are called "axial" because the piston cylinder block is varied from parallel to the axis of the drive shaft to several degrees out of parallel, where flow is the greatest. The cylinder block and thrust plate rotate together, as indicated by Figure 13. When the pistons move away from center, they are filled with fluid from one side of the pintle. When they compress toward the center, they expel fluid at the outlet on the other side. The axial piston pump is the most popular piston pump, since it is more compact and less complex than the radial type.

The most common use of positive displacement pumps is in hydraulics, where they are used to pressurize hydraulic fluid.

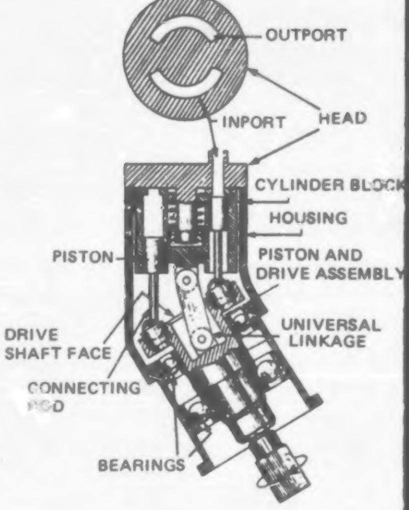

OUTPORT
INPORT
HEAD
CYLINDER BLOCK
HOUSING
PISTON
PISTON AND DRIVE ASSEMBLY
UNIVERSAL LINKAGE
DRIVE SHAFT FACE
CONNECTING ROD
BEARINGS

FIG. 12

SUMMARY

Fluid pumps convert mechanical energy into fluid energy, and are used to circulate or to pressurize fluid. Hydrodynamic pumps are used primarily to circulate fluids. Positive displacement pumps are used in pressure applications often required in hydraulic systems. The four characteristic differences between these two pump types are defined in Table 1.

Pump characteristic curves usually are furnished by pump manufacturers for each size pump. These curves are used to determine the size, or capacity, of pumps required for given applications.

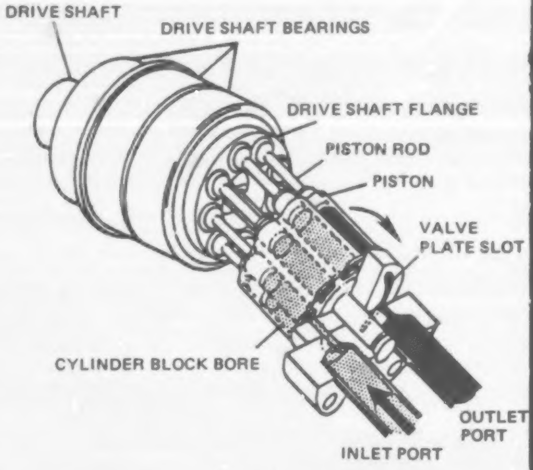

DRIVE SHAFT
DRIVE SHAFT BEARINGS
DRIVE SHAFT FLANGE
PISTON ROD
PISTON
VALVE PLATE SLOT
CYLINDER BLOCK BORE
OUTLET PORT
INLET PORT

FIG. 13

TABLE 1. CHARACTERISTIC DIFFERENCES BETWEEN HYDRODYNAMIC AND POSITIVE DISPLACEMENT PUMPS

Characteristic	Hydrodynamic	Positive Displacement
Pressure	Low	High
Volume Flow Rate	High and variable	Low and constant
Priming	Required	Not required
Danger of blocked output	Slippage prevents damage	Damage to pump

PREPARATION OF PUMP CHARACTERISTICS CURVE

MATERIALS

Small vertical a.c. centrifugal pump with output of approximately 5 gal per min.
Two water containers: one for pump at approximately 5 gal capacity and one for use as receptacle at approximately 2 gal capacity, marked at 1 gal point.
a.c. Wattmeter.
Laboratory timer.
Yardstick or tape measure.

PROCEDURE

1. Connect pump, hose, and containers as illustrated in Figure 14.

2. Connect a.c. Wattmeter to pump; turn on power to pump.

3. Measure power in watts consumed by pump motor in pumping one gallon of water with hose connected for head of 0.5 ft from water level. Also measure time required to pump one gallon. Record watts and time in Data Table.

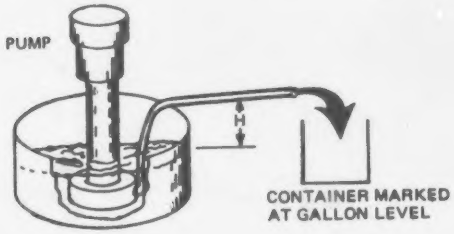


FIG. 14

4. Repeat step 3 for the following heads. 1.0, 1.5, 2.0, and 2.5. Record.

5. Calculate and record pressure at each head. Use Equation 1.

$$p = \rho_w h \tag{1}$$

where
ρ_w = weight density of water (62.4 lb/ft^3).
h = head (height) in feet.

6. Calculate flow rate in gal/min for each pressure head. Record.

7. Calculate and record output energy for each head. Use Equation 2.

$$E_{out} = \rho_w hV \qquad (2)$$

where
ρ_w = weight density of water (8.34 lb/gal).
V = volume in gallons.

8. Calculate and record power output in ft·lb/s for each head. Use Equation 3.

$$P_{out} = \frac{E_{out}}{t} = \frac{\rho_w hV}{t} = (\rho_w h)\left(\frac{V}{t}\right) \qquad (3)$$

or

Power = pressure x volume flow rate

9. Calculate output horsepower for each head by dividing the P_{out} value obtained in Equation 3 by 550 (since 1 hp = 550 ft·lb/s). Record.

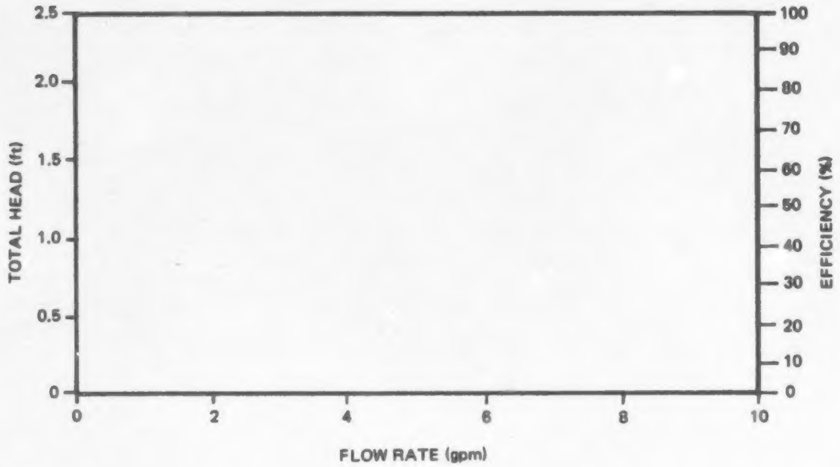
10. Calculate electrical output horsepower by dividing the wattage measured at each head by 746 (since 746W = 1 hp). Record.

11. Calculate and record % efficiency of pump at each head, using values obtained and Equation 4.

$$\eta = \frac{E_{out}}{E_{in}}(100) \qquad (4)$$

12. Transfer head and overall efficiency data from Data Table to Figure 15, and plot curves.

13. Compare the curves plotted in Figure 15 to those in Figure 3 and Example A.



DATA TABLE

Head (ft)	Pressure lb/in^2	P_{in} W	Time (s)	Flow Rate gal/min	E_{out} ft·lb	P_{out} ft·lb/s	P_{out} hp	P_{out} W	% Efficiency η
0.5									
1.0									
1.5									
2.0									
2.5									

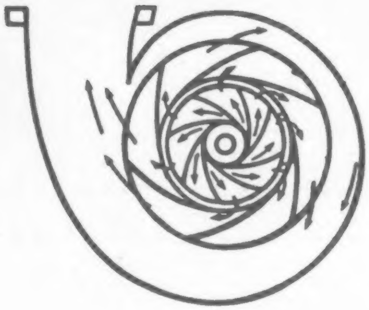
SELF CHECK

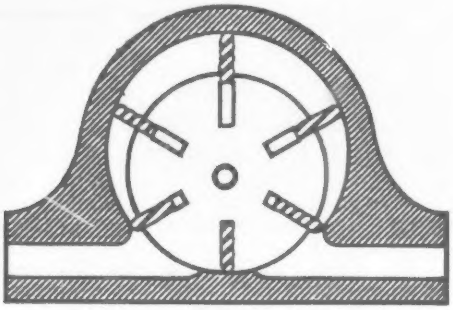
1. What are four characteristic differences between hydrodynamic pumps and positive displacement pumps?

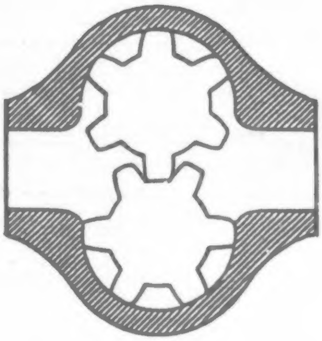
2. What is the most common application of the hydrodynamic pump?

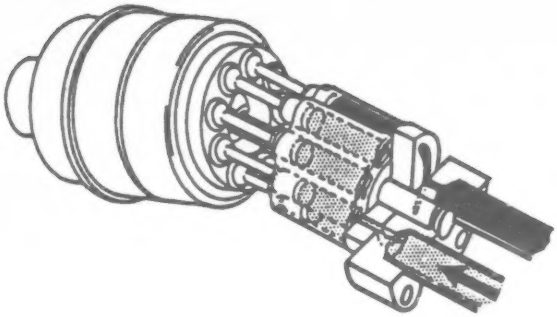
3. What is the most common application of the positive displacement pump?

4. Identify and label the pumps in the following diagrams. Label the parts, their direction of movement, and direction of fluid flow.





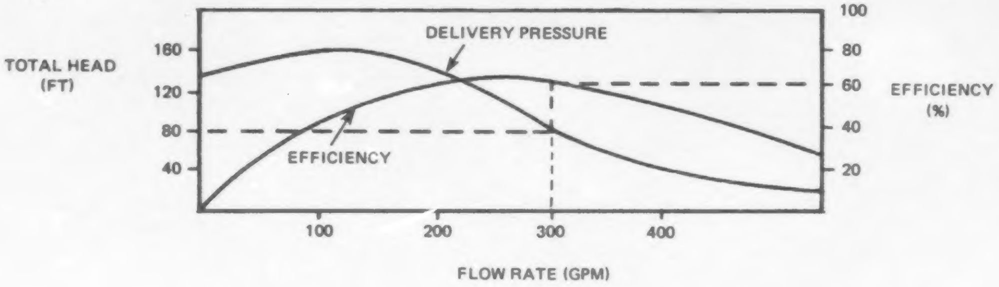




5. How can unbalanced vane pumps and radial and axial piston pumps be used as variable-volume pumps?

6. How can a pump characteristic curve be of assistance in determining the size of pump required for a specific application?

7. Find the flow rate and % efficient of a pump with the characteristic graph shown below when there is a total head of 80 ft.



8. What is the pressure at a head which has a height of 2 ft if the fluid used is water?

9. What is the head height in a pump which has an output energy of 16 ft·lb, if the volume of water pumped is 1 gallon?

10. What is the efficiency of a pump with an input power of 4W, a pressure head height of 2 ft, and a volume flow rate of 3 gallons per minute?

11. The operating voltage of an electrical pump is 110 volts. If the pump has an energy output of 20 joules every 20 seconds and is 40% efficient, what current does the pump draw?

FANS AND BLOWERS

Fans and blowers are energy converters that use both electrical and mechanical energy to produce fluid energy. In a fan or blower, the electrical energy of the motor is converted into the rotational mechanical energy of the fan blades. The mechanical energy of the fan blades then moves the air through the discharge duct, thereby creating fluid energy. The ability to move air and other fluids through ducting systems is basic to air conditioning and heating technologies and plays a major role in many other areas of technological society. This module presents the theory of operation and basic design of fans and blowers. In the laboratory, the student will determine the air horsepower and the efficiency of a fan motor.

OBJECTIVES

Upon completion of this module, the student should be able to:

- Define the following terms:

  Air horsepower.
  Power input.
  Static efficiency.
  Static pressure.
  Total pressure
  Velocity pressure
  Mechanical efficiency.

- Given the static pressure, velocity pressure, volume flow rate, and shaft horsepower of a fan, determine its static and mechanical efficiency.

- Given a schematic representation of an axial flow fan and a radial flow fan, correctly identify each; and indicate the direction of air flow on both diagrams.
- Given the schematic representation of six different blade types in a centrifugal (radial) fan, correctly identify and label each of the blade types.
- Given a schematic representation showing four different inlet designs for a propellor, draw the air flow for each design. Indicate which of these four systems is most efficient.
- Given the appropriate equipment, measure the velocity of air through an air discharge duct and the static and velocity pressures; from these measurements determine the air horsepower and the efficiency of the fan motor.

DISCUSSION

Fans, blowers, and compressors move air, but at greatly different pressures. Fan pressures range from a few inches of water up to about 1-psi. Blowers work to about 50-psi; compressors operate from about 35-psi and up. Since fans, blowers, and compressors all basically do the same thing, this module discusses only fans. When appropriate, important differences between the three will be pointed out.

A fan can be defined as a volumetric machine that operates very much like a pump. In simple terms, a fan moves quantities of gas or air from one place to another. In doing this, it overcomes resistance to flow by supplying the fluid with the energy necessary for continued motion. Physically, the essential elements of a fan are a bladed rotor (such as an impeller or propeller) and a housing in which to collect the incoming air or gas and to direct its flow. Figure 1 displays the essential elements of a fan.

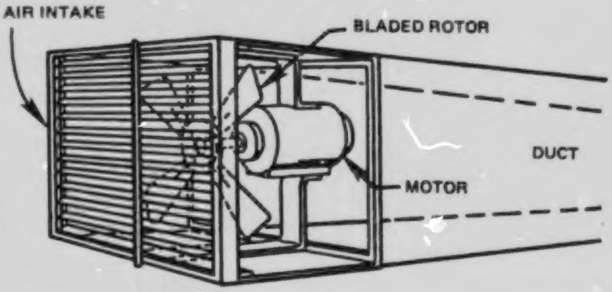


FIG. 1

WORK OF A FAN

Since any device that makes something else move is doing work, a fan requires energy for operation. The amount of energy required depends upon the volume of gas moved, the resistance against which the fan works, and the machine efficiency.

These energy relationships are clarified by the pressure-volume diagram shown in Figure 2 for a typical fan and compressor. From point 1 to point 2 in the fan cycle, there is very little pressure change in the suction phase. During this phase, air is sucked into the intake of the fan at constant pressure until the volume at point 2 is reached. At this point the fan blade compresses the air slightly, increasing the pressure. This step is shown as the compression phase in Figure 2 (point 2 to point 3). From point 3 to point 4, the fan is discharging the gas at constant pressure. At point 4, the gas has been completely discharged from the fan

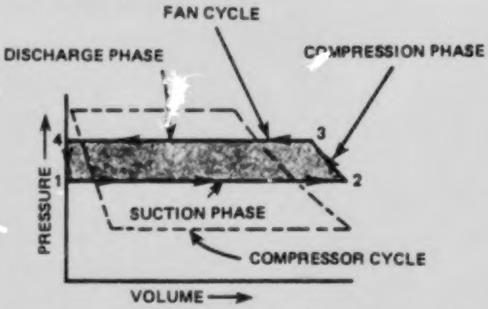


FIG. 2

outlet. The fan then shows a drop in pressure back to point 1, where the cycle is repeated. The work done by the fan is equal to the area enclosed by the fan cycle. Figure 2 also demonstrates that the pressure changes involved are relatively minor.

Also shown in Figure 2 is the pressure-volume diagram for a compressor. Notice that the compressor cycle operates over a much larger pressure range; otherwise, the diagram is similar in every respect to that for the fan. Unlike blowers or compressors, which work against larger pressures and significant pressure increases, the bulk of fan work goes into moving gas at a relatively low pressure. Notice also from Figure 2 that the volume remains virtually constant during the compression phase because the change in specific weight of the gas between fan inlet and discharge is negligible; in fact, for a 1-psi pressure increase, the specific weight of air changes by less than 7%. Thus, when fan power is analyzed, the volume change during the compression phase can be thought of as zero. Work accomplished by the fan cycle is expressed accurately by Equation 1.

$$W = V\Delta p \tag{1}$$

where

W = work done in fan cycle.

Δp = rise in pressure across the fan.

V = volume of air or gas moved.

Pressure difference (either static or total) usually is measured in inches of water. Static pressure is the force per unit area exerted on walls and ducts — like the steam pressure on the inside of a boiler drum. Static pressure is that pressure that overcomes the resistance of the ducts.

Total pressure represents the combined effect of static pressure and velocity pressure. Velocity pressure is that pressure over and above static pressure caused by the movement of the air or gas. Figure 3 illustrates how manometers inserted in a fan discharge duct measure static pressure, velocity pressure, and total pressure. The total pressure in a duct is given by Equation 2.

$$p_t = p_s + p_v \tag{2}$$

where

p_t = total pressure in duct.

p_s = static pressure.

p_v = velocity pressure.

AIR HORSEPOWER

From Equation 1, an equation can be developed to express "air horsepower." Air horsepower is defined as "the horsepower required at 100% fan efficiency to move a given quantity of air against a given pressure." The resulting expression for air horsepower Ahp is given by Equation 3.

$$\text{Ahp} = \frac{Q_V \Delta p}{6356} \qquad (3)$$

where

Ahp = air horsepower.

Q_V = volumetric flow (flow rate) through the fan in cubic feet per minute.

Δp = pressure difference across the fan in inches of water.

The air horsepower also may be designated as "static" or "total." Since the resistance to be overcome in fan applications is primarily static pressure, the fan pressure developed usually is referred to as "static pressure." On this basis, the calculated fan power is known as "static air horsepower," as given by Equation 4.

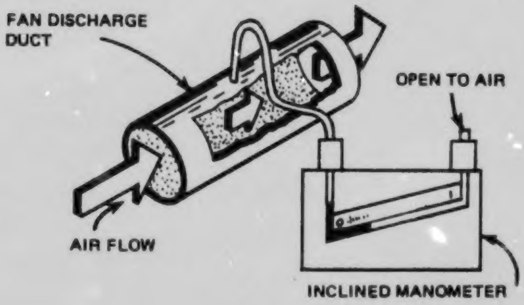
$$\text{Ahp}_s = \frac{Q_V \Delta p_s}{6356} \qquad (4)$$

where

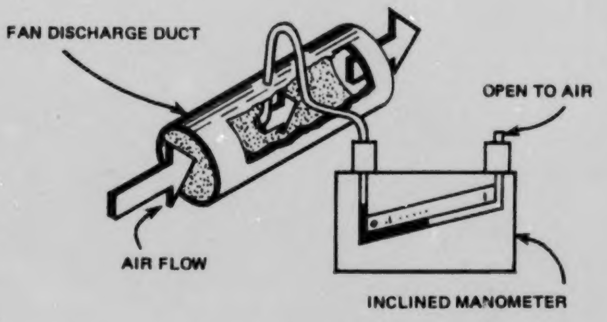
Ahp_s = static air horsepower.

Δp_s = static head pressure in inches of water.

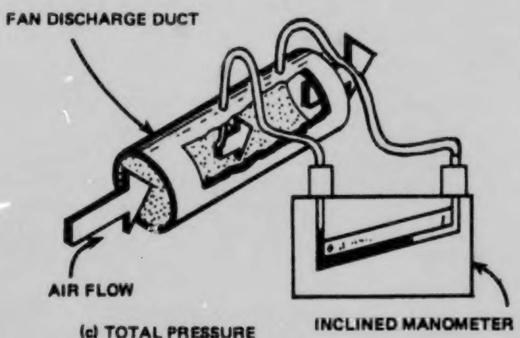
When the power calculations are based upon total pressure, fan power is referred to as the "total air horsepower" and is equivalent to the output


FAN DISCHARGE DUCT
OPEN TO AIR
AIR FLOW
INCLINED MANOMETER

(a) STATIC PRESSURE


FAN DISCHARGE DUCT
OPEN TO AIR
AIR FLOW
INCLINED MANOMETER

(b) VELOCITY PRESSURE


FAN DISCHARGE DUCT
AIR FLOW
INCLINED MANOMETER

(c) TOTAL PRESSURE

FIG 13

power. Total air horsepower is derived from Equation 5.

$$\mathrm{Ahp}_t = \frac{Q_v \Delta p_t}{6356} \tag{5}$$

where

Ahp_t = total air horsepower.
Δp_t = total head pressure in inches of water.

EFFICIENCY

In similar fashion, static efficiency is associated with Ahp_s; and mechanical efficiency, or total efficiency, is associated with power output. Each is compared with shaft horsepower, or the power input to the fan, to arrive at an actual efficiency figure. Equation 6 is for static efficiency.

$$\eta_s = \frac{\mathrm{Ahp}_s}{P_{in}} \tag{6}$$

where

η_s = static efficiency.
P_{in} = power input to fan.

Mechanical, or total, efficiency is given by Equation 7.

$$\eta_t = \frac{\mathrm{Ahp}_t}{P_{in}} \tag{7}$$

where

η_t = mechanical (total) efficiency.

Many small variables influence the power output efficiency of any fan; such things as pressure and temperature are the most important. For the purposes of this module, no corrections are made for temperature and pressure variations of the air or gas.

Equations 6 and 7 are utilized to find the solution to a problem in Example A.

EXAMPLE A: STATIC AND MECHANICAL EFFICIENCY

Given: A fan used for air conditioning and heating develops 4.5-inch static pressure and 0.7-inch velocity pressure when flow is 8,000 cfm of 100°F air; shaft horsepower is 7.2 hp.

Find: (a) Static efficiency.

(b) Mechanical efficiency.

Solution: (a) $\text{Ahp}_s = \dfrac{Q_V \Delta p_s}{6356}$ (Equation 4)

$$= \frac{(8000)(4.5)}{6356}$$

$$\text{Ahp}_s = 5.7 \text{ hp}$$

$\eta_s = \dfrac{\text{Ahp}_s}{P_{in}}$ (Equation 6)

$$= \frac{5.7 \text{ hp}}{7.2 \text{ hp}}$$

$$\eta_s = 79\%$$

(b) $\Delta p_t = \Delta p_s + \Delta p_v$ (from Equation 2)

$$= 4.5 \text{ in} + 0.7 \text{ in}$$

$$\Delta p_t = 5.2 \text{ in}$$

$\eta_t = \dfrac{\text{Ahp}_t}{P_{in}}$ (Equation 7)

$$= \frac{6.54}{7.2}$$

$$\eta_t = 91\%$$

THE TWO FAN TYPES

There are two basic classes of fans: (1) the axial-flow design, which moves gas or air parallel to the axis of rotation, and (2) the centrifugal-flow (or radial-flow) type, which moves air or gas perpendicular to the axis of rotation. Axial-flow fans are better suited to low resistance applications. Centrifugal-flow fans usually take care of the higher pressure jobs.

The axial-flow fan uses the screwlike action of a multi-bladed rotating shaft or propeller to move air or gas in a straight-through path. The leading edge of each rotating propeller blade bites into the air or gas, which is then propelled through the fan and discharged in a helical pattern, as shown in Figure 4, by the blades' trailing edge.

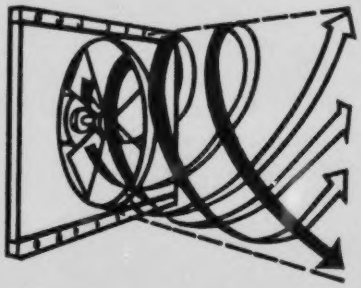
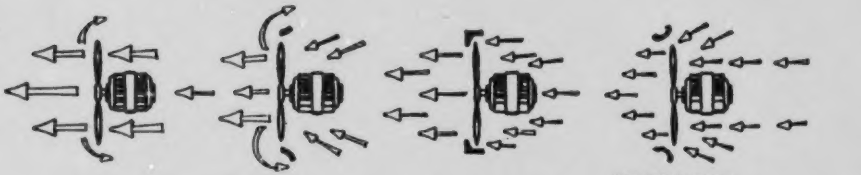


FIG. 4

The efficiency of a propeller fan depends largely upon its inlet design because air recirculation around blade tips greatly reduces the efficiency of a propeller-type fan. Installation of a curved orifice-like ring at the blade tips reduces the amount of recirculation and improves efficiency. An angle-shaped ring almost eliminates recirculation. Optimum efficiency is achieved with a combination curved, angle-shaped ring, which also smoothens flow. The flow patterns for the air recirculating around the blade tips of these various inlet designs is illustrated in Figure 5.



NO RING CURVED ORIFICE ANGULAR ORIFICE COMBINATION CURVED - ANGULAR ORIFICE

FIG. 5

Two common varieties of axial flow fans are exhibited in Figure 6.

The tube-axial design is nothing more than a propeller fan enclosed in a cylinder that collects and directs air flow. In a vane-axial fan (an extension of the tube-axial design), air guide vanes on the air discharge side of the propeller straighten out the air flow pattern and increase static pressure.

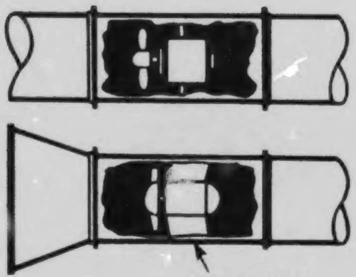


FIG. 6

Axial-flow fans can have widely-differing characteristics, depending upon the design, blade type, and ducting. Generally, they are applied where the required outlet velocities are higher than can be produced by centrifugal-fans and when pressure demands are not above actual flow limits.

Centrifugal, or radial, fans are advantageous when the air must be moved in a system in which the frictional resistance is relatively high (Figure 7). The air drawn into the center of the revolving wheel turns 90° and enters the space between the blades in the wheel. The bladed wheel whirls air centrifugally between each pair of blades and tosses it out peripherally at high velocity and increased static pressure. As this happens, more air is sucked in at the eye of the wheel. As the air leaves the revolving blade tips, part of its velocity is converted into additional static pressure by divergence of the scroll-shaped housing.

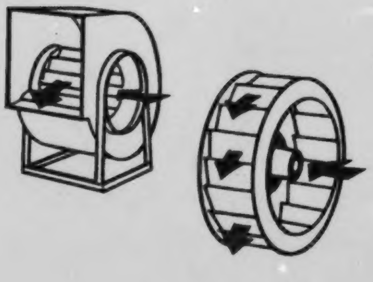


FIG. 7

Blade shapes employed are of three basic types: forward curved, straight, and backward curved. Other configurations, including an air foil design, are merely variations of these designs. Some commonly-used shapes are given in Figure 8.

In general, blade type limits top fan speed. Backward-curved blade machines can operate at a higher speed than forward-curved designs. The type of blade employed also depends upon space limitations, allowable noise levels, efficiency demanded by specified load conditions, and desired fan performance characteristics. While mechanical efficiencies are much the same for each different blade type, differences in horsepower and pressure-volume relationships can be significant.

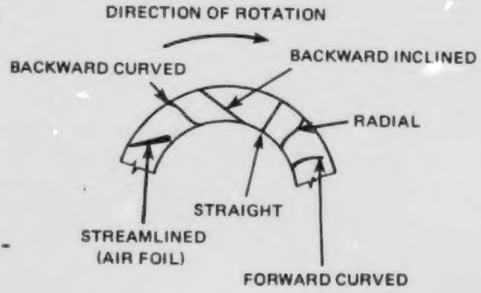


FIG. 8

SUMMARY

A fan requires energy to move gas or air through a duct system. The amount of energy required depends upon the volume of gas moved, the resistance of the duct system, and the machine efficiency. The efficiency of a fan can be calculated once the volumetric flow through the fan and the pressure difference across the fan are known. There are two basic classes of fans: the axial-flow design and the centrifugal-flow type. Each of these classes has efficiencies that depend, amoung other things, upon inlet design and blade types. In general, axial-flow fans are better suited to low-resistance applications; and centrifugal flow fans are used in higher-pressure applications.

AIR HORSEPOWER AND EFFICIENCY OF A FAN MOTOR

MATERIALS

Fan — Axial fan with electric motor. 110-120 V, 60 cycle, 0.6 A

Manometer — Dwyer incline portable Module 102 or Ellison inclined portable with hemispherical pitot.

PROCEDURE

1. Arrange fan and discharge duct as shown in Figure 9. If fan does not have a discharge duct, fabricate one out of galvanized steel or other suitable material.

2. Record fan horsepower in Data Table as power input.

3. With fan on, measure the total pressure, static pressure, and velocity pressure as shown in Figure 3. Record results in Data Table.

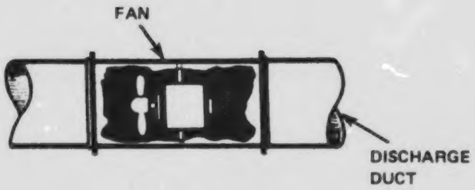


FIG. 9

4. Calculate air velocity from the relationship —

$$v = \frac{1096.2\ p_v}{\rho}$$

where

v = air velocity.
p_v = velocity pressure in inches of water.
ρ = air density (0.075 lb/cu ft).

Record in Data Table.

5. From the following relationship, calculate volumetric flow rate; record in Data Table.

$$Q_V = Av$$

where

Q_V = volumetric flow rate in ft^3/min.
A = duct area in ft^2.
v = air velocity in ft/min.

6. Calculate and record the static air horsepower Ahp_t and static efficiency η_s. Record results in Data Table.

$$Ahp = \frac{Q_V \Delta p_s}{6356}$$

$$\eta_s = \frac{Ahp_s}{P_{in}}$$

7. Calculate and record the total air horsepower Ahp_t and the mechanical efficiency η_t.

$$Ahp_t = \frac{Q_V \Delta p_t}{6356}$$

$$\eta_t = \frac{Ahp_t}{P_{in}}$$

DATA TABLE

Power Input = P_{in} = ______________________ hp

Total Pressure = $p_s + p_v$ = ______________________ in of H_2O

Static Pressure = p_s = ______________________ in of H_2O

Velocity Pressure = p_v = ______________________ in of H_2O

Air velocity = v = ______________________ ft/min

Flow rate = Q_V = ______________________ ft^3/min

Ahp_s = ______________________ hp

Static Efficiency (η_s) = ______________________ %

Ahp_t = ______________________ hp

Mechanical Efficiency (η_t) = ______________________ %

SELF CHECK

1. Write a formula for the following terms:

Total pressure p_t =

Static air horsepower Ahp_s =

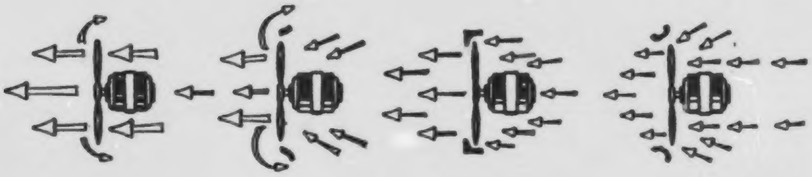
Total air horsepower Ahp_t =

Static efficiency η_s =

Mechanical efficiency η_t =

2. Prepare a drawing that illustrates the air flow through a radial-flow fan and through an axial-flow fan.

3. Label the blade types illustrated below, and identify the most efficient fan inlet shown.



4. A fan produces a static pressure of 2.5 inches of water and a velocity pressure of 0.5 inches of water. The flow rate is 3600 cfm. The input horsepower is 2 hp. Determine the static efficiency and the mechanical efficiency of the fan.

5. A fan with an input power of 1 hp produces a flow rate of 2000 cfm with a static pressure of 2.6 inches of water. The mechanical efficiency of the fan exceeds the static efficiency by 10%. Determine the velocity pressure produced.

6. How much work does a fan do in one cycle if the pressure rise across the fan is 0.02 N/m^2 and the volume of air moved is 0.1 m^3?

7. How much power must be added to a fan in order to achieve an air horsepower of 7 hp if the mechanical efficiency of the fan is 72?

8. What is the total air horsepower of a fan if the total head pressure is 5 inches of water, the velocity pressure is 1 inch of water and the duct area is 2 ft^3? (Hint: See Laboratory Section.)

9F5

VACUUM PUMPS

VACUUM PUMPS

A number of different pumps are used to produce a vacuum in enclosed chambers. These pumps include rotary mechanical pumps, diffusion pumps, sputter-ion pumps, and cryogenic pumps. The pressure ranges that can be covered vary from one pump type to another. In common practice, a mechanical pump is used to pump down from atmospheric pressure to medium vacuum; and then a second pump is switched on to reach high or ultrahigh vacuum.

In the laboratory, the student will use at least two different pump types to evacuate a vacuum system and will measure the pressure in the system during pump-down.

OBJECTIVES

Upon completion of this module, the student should be able to:

- Define the following terms:

  Pumping speed.
  Conductance for flow of gas through a pipe.
  Impedance for flow of gas through a pipe.
  Pump-down time.

- Given the necessary information, calculate the time required for a pump to pump down a specific volume.

- Given the necessary information, calculate the impedance for flow of gas through a specific section of pipe.

- Describe the operation of the following types of pump:

  Rotary mechanical pumps.
  Diffusion pumps.
  Sputter-ion pumps.
  Cryogenic pumps.

- State the approximate pressure range covered by each of the above pump types.

- Given the appropriate equipment, evacuate a vacuum chamber, and measure the pressure during pump-down, using —

  A mechanical pump alone.
  A mechanical pump plus either an oil-diffusion pump or a sputter-ion pump.

DISCUSSION

This module emphasizes the operation of the various types of pumps commonly used to evacuate vacuum systems. The concepts of degree of vacuum, gas pressure, and mean free path presented in Application Module 10F5 will be taken for granted.

A pump for reducing pressure in a chamber may be characterized by its pumping speed. The pumping speed measures the rate at which gas is removed by the pump. It has units of volume per unit time; for example, liter/second. The definition of pump speed is the volume of gas per unit of time removed from the system at the pressure existing at the inlet to the pump. Note: Pump speed is specified for the pressure at the inlet. As the pressure is decreased for a constant pump speed, therefore, the volume of gas removed is the same; but since the pressure of the gas in continually being reduced, the quantity of gas (e.g., the number of gas molecules) removed per unit time is continually decreasing.

Pump speed is a function of gas pressure. Figure 1 shows a typical curve that defines pump speed versus pressure. At some high value of pressure, the pump will not work well, and the pump speed will be low. Also, when the pressure becomes very low, the pump again is incapable of removing the gas. Over a range of pressures between these limits, the pumping speed will be reasonably constant. (The pressure range over which various types of pumps are useful will be described later.) The pressure range over which pump speed is constant may cover several orders of magnitude.

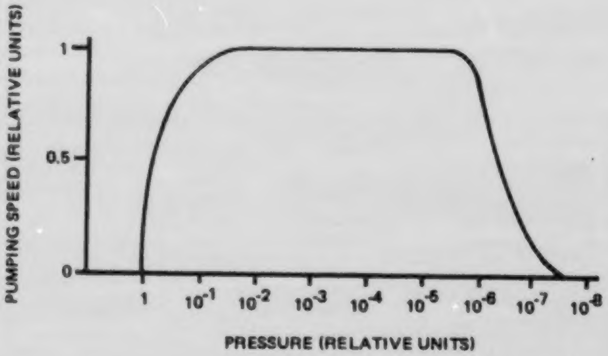


FIG. 1

The time required to evacuate a given system from initial pressure to final pressure is related to the pumping speed by Equation 1.

$$t = \frac{2.3\ V}{S} \log_{10} \frac{p_1}{p_2} \tag{1}$$

where

t = time required to evacuate system.
V = volume of system.
S = pumping speed.
p_1 = initial pressure.
p_2 = final pressure.

Equation 1 holds true if the initial and final pressures lie within the range at which the pump is designed to operate. Equation 1 is used in Example A to calculate pumping speed.

EXAMPLE A: CALCULATION OF PUMPING SPEED

Given: A mechanical pump operates at a speed of 100 liters/s, over the pressure range from 760 torr to 1 torr.

Find: Time required for this pump to evacuate a 50-liter chamber from atmospheric pressure to 1 torr.

Solution: Use Equation 1:

$$t = \frac{2.3\ V}{S} \log_{10} \frac{p_1}{p_2}$$

$$= \frac{(2.3)(50)}{100} \log_{10} \frac{760}{1}$$

t = 3.3 seconds

In practice, pumps are connected to the vacuum chamber through pipes or other channels, through which the gas must flow. The pipe also will affect the rate of gas removal. The effect of the pump is expressed in terms of a parameter called "conductance." Conductance for a channel is defined as "the flow of gas molecules through the channel per unit time, divided by the pressure difference between the ends of the channel." Mathematically, conductance can be expressed by Equation 2.

$$C = \frac{p_1 V_1}{(p_1 - p_2)} \qquad (2)$$

where

C = conductance in volume/time.

V_1 = volume of gas (at initial, inlet pressure p_1) that moves through the channel per unit time.

The product p_1V_1 is, from the ideal gas law, a measure of the quantity of gas so moved, per second.

The reciprocal of the conductance is called the "impedance" and is given by Equation 3.

$$Z = \frac{1}{C} \qquad (3)$$

where

Z = impedance.

Notice the analogy between Equations 2 and 3 and resistance and its reciprocal in the electrical system. Thus, two pipes with impedance Z_1 and Z_2 connected in series have total impedance Z_T indicated by Equation 4.

$$Z_T = Z_1 + Z_2 \tag{4}$$

For the same two pipes connected in parallel, the total impedance Z_T is given by Equation 5.

$$\frac{1}{Z_T} = \frac{1}{Z_1} + \frac{1}{Z_2} \tag{5}$$

Equations 4 and 5 are exactly similar to the equations for the electrical resistance of two resistors.

Determination of conductance in a practical case can become complicated. The exact value of conductance depends upon the shape of the pipe, the average pressure, and the dimensions of the pipe, in a complicated fashion. Only one valid relation in relatively low pressure regions (pressure $<10^{-2}$ torr) is presented here; in this region, for a circular pipe, Equation 6 applies.

$$C \simeq \frac{12\ d^3}{\ell} \tag{6}$$

where

C = conductance in liters/s.

d = diameter of pipe in cm.

ℓ = length of pipe in cm.

The effect of the pipe's conductance C on the pumping speed of the total system is given by Equation 7.

$$\frac{1}{S_T} = \frac{1}{S} + \frac{1}{C} \tag{7}$$

where

S_T = pumping speed of total system (pump plus pipe).

S = pumping speed of pump alone.

This equation implies that the pumping speed of the system cannot be greater than the speed of the pump or of the conductance of any of the pipes in the system. Equation 7 is used to solve a problem in Example B.

EXAMPLE B: CALCULATION OF PUMPING SPEED

Given: The pump of Example A is connected to the chamber of volume 50 liters by a pipe of 40 liters/second conductance.

Find: Time required to pump the chamber down from 760 torr to 1 torr.

Solution: The total pumping speed S_T is given by Equation 7:

$$\frac{1}{S_T} = \frac{1}{100} + \frac{1}{40} = 0.035$$

$$S_T = 28.6 \text{ liters/s}$$

When this new total pumping speed is substituted into Equation 1, the result is the following:

$$t = \frac{2.3\ V}{S} \log_{10} \frac{p_1}{p_2}$$

$$= \frac{2.3 \times 50}{28.6} \log_{10} \frac{760}{1}$$

$$t = 11.6 \text{ seconds}$$

TYPES OF VACUUM PUMPS

Many different types of vacuum pumps are in use, including these four common types:

- The rotary mechanical pump.
- The diffusion pump.
- The sputter-ion pump.
- The cryogenic pump.

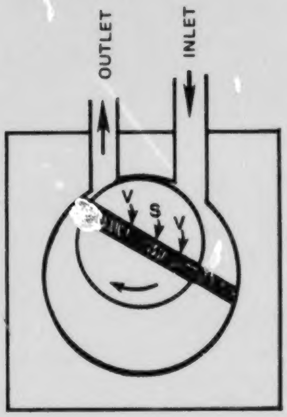
Each type is designed for a range of pressures for which it is effective. Only the rotary mechanical pump exhibits a range that includes atmospheric pressure. Thus, a common situation involves use of a mechanical pump plus a second pump; for example, a diffusion pump. The mechanical pump reduces the pressure from 760 torr down to the lower limit of its operating range, usually in the middle-vacuum range. Then the diffusion pump is switched on. It is effective in the middle-vacuum range and pumps the system down to a lower final pressure.

ROTARY PUMPS

Rotary mechanical pumps rely upon the motion of internal structures in the pump. In such pumps, during a cycle of its rotation, gas from the vacuum system is trapped, compressed, and moved from the lower-pressure inlet to the high-pressure outlet. Gas then is expelled to the atmosphere.

Figure 2 is a diagram of a typical rotary mechanical pump. This design is a common one employed in many commercial pumps, particularly in smaller pumps. There are other designs for mechanical pumps, but the basic features of mechanical pumps can be understood with reference to Figure 2.

The circular rotating element, or rotor, is driven by an electric motor. The rotor is inside a larger circular area within the pump body. The rotating element is off center, as shown, so that one side always is in contact with the interior wall. The empty volume is crescent-shaped, as shown. The rotating element has two vanes that are forced outward by a spring. The vanes maintain contact with the walls throughout the rotation of the rotor.



V - VANE

S - SPRING

FIG. 2

Oil lubricates the rubbing surfaces and seals the interfaces, in order that gas cannot flow past the ends of the vanes.

Gas enters the pump from the chamber through the inlet. Figure 3 illustrates stages in the rotation cycle of the rotor. In Figure 3a, gas enters the pump. In Figure 3b, the vane has passed the inlet and the gas is trapped. The volume available to the trapped gas is reduced, and the gas is compressed (Figure 3c). In Figure 3d, the volume has become very small, and the gas is expelled throughout the outlet. Meanwhile, more gas has entered through the inlet for another cycle.

Rotary mechanical pumps are effective over a pressure range of from 760 torr down to a lower limit of about 5×10^{-3} torr. Thus, they can be used as the first stage in a multi-pump system. The mechanical pump reduces the pressure from atmospheric pressure down to the lower limit of its operating range. Then a second

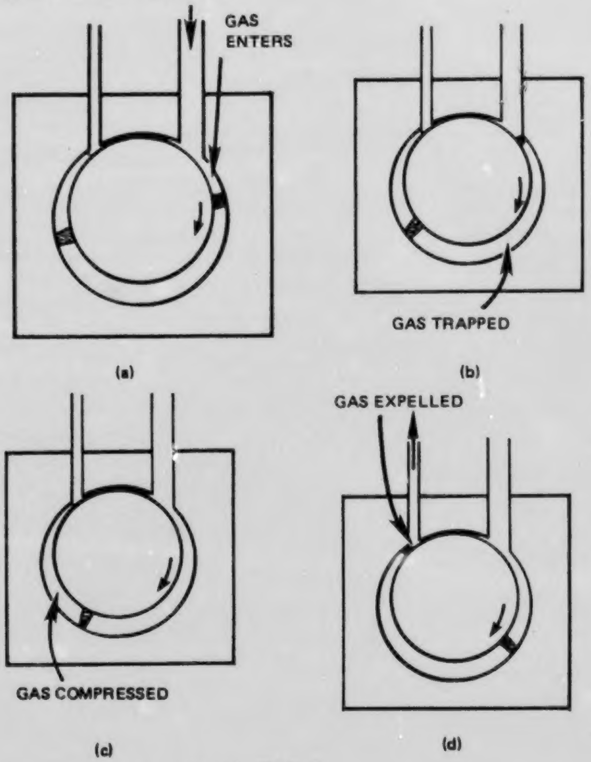


FIG. 3

pump, capable of reaching lower pressures, is switched on and further reduces the pressure.

The pumping speeds of available mechanical rotary pumps vary over a wide range, from small units having speeds of less than 1 liter/second, driven by 1/8-horsepower motors, up to units with pumping speeds of over 300 liter/second, driven by 40-horsepower motors.

DIFFUSION PUMPS

Diffusion pumps remove gas by forming a high velocity jet of vapor from oil or mercury evaporated in the low pressure produced by a mechanical pump. The high-velocity jet carries the gas molecules along with it to the discharge port that leads to the mechanical pump.

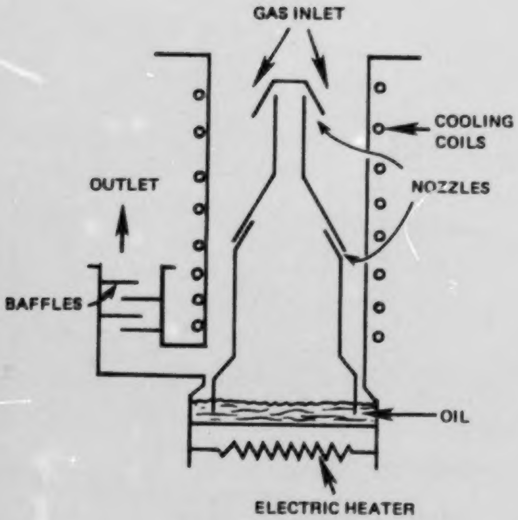


FIG. 4

A diffusion pump is illustrated in Figure 4. The outlet is connected to a mechanical pump that exhausts the pumped gas to the atmosphere. The pump shown is an oil diffusion pump, but mercury may be used instead. The oil is evaporated by the heater. It passes upward in the central part of the pump and emerges through the nozzles. As the oil vapor passes through the nozzles, it is deflected downward and carries gas molecules that enter through the inlet along with it. The oil is recovered by condensation on the pump walls, usually cooled by water running through coils wrapped around the pump. The gas emerges from the outlet, where baffles trap any remaining oil vapor. Figure 4 displays two stages of nozzles, but diffusion pumps with more stages are common.

Diffusion pumps require a pressure below one atmosphere to begin operation; therefore, they always are used in conjunction with a mechanical pump. The mechanical pump reduces the pressure to about 0.1 torr or so, at which pressure the diffusion pump can begin operation. The diffusion pump then can reduce the pressure much further. The lower limit of operating pressure for diffusion pumps is in the range from 10^{-6} to 10^{-9} torr, depending upon the design.

The pumping speeds offered by diffusion pumps cover a wide range, depending upon size and type of construction. Diffusion pumps with speeds ranging from 1 liter/second to many thousands of liters/second are available.

All diffusion pumps suffer, to some extent, from a backstreaming of the oil vapor into the vacuum chamber, which leads to contamination of the chamber by the oil. To minimize this, cold traps and baffles are placed between the pump and the chamber. Such baffles increase the probability that an oil molecule will strike a cold surface and condense before making its way into the vacuum chamber.

SPUTTER-ION PUMPS

In sputter-ion pumps, the gas is removed from the vacuum chamber volume by being trapped on the surfaces of electrodes within the pump. The gas is not expelled, but retained within the pump. The gas molecules are trapped, both by chemical reactions with a reactive metal that forms the electrodes, and also by being buried under newly-sputtered layers of metal. The metal most commonly used for sputter-ion pumps is titanium, which is highly active chemically.

A schematic diagram of the structure of a sputter-ion pump is shown in Figure 5. A voltage of several kilovolts is applied between the anode and the cathode, supporting an electrical discharge in the low-pressure gas. Positive ions of the gas molecules are produced in the discharge. These ions are accelerated to the cathode; they strike the cathode with high velocity and are buried within the cathode. The impact of the high-velocity ions drives titanium atoms from the surface of the cathode; that is, it "sputters" titanium atoms. The reactive titanium atoms react chemically with gas molecules and then settle on the walls of the pump. The gas molecules, thus, are effectively removed. The magnet of Figure 5 causes the electrons in the electrical discharge to move in spiral paths and increases the probability of ionization for gas molecules. The magnetic field typically is about 3000 Gauss.

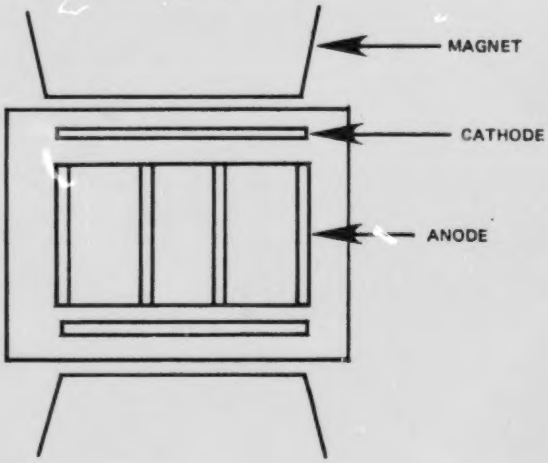


FIG. 5

The titanium ion pump is very effective for chemically-active gases, such as oxygen, nitrogen, and hydrogen. It is less effective for inert gases, especially for argon, which is removed only with difficulty by such a pump.

Sputter-ion pumps are desirable when a system free from oil vapor is desired. They are entirely free from such contamination.

Because the trapped gas builds up on the pump walls, the total capacity of such a pump is limited. Sputter-ion pumps must be cleaned periodically to remove the gas entrapped on the walls. In order to limit the total amount of gas that must be trapped, they are started only at low pressure, 10^{-2} torr or less. They, too, also are used in conjunction with a mechanical pump, which reduces the pressure to the range at which the sputter-ion pump is effective. Then the sputter-ion pump is started. Final gas pressures as low as 10^{-10} torr can be obtained with available pumps of this type. Sputter-ion pumps with pumping speeds ranging from 10 liters/second to 10^4 liters/second are available. They offer an effective and clean method of reaching the ultrahigh vacuum region.

CRYOGENIC PUMPS

In cryogenic pumps, a surface is held at very low temperature by contact with liquid nitrogen (temperature 77 K) or liquid helium (temperature 4 K). Other gases striking the surface will be condensed and trapped on it. As is the case with sputter-ion pumps, the gas is not expelled from the system, but trapped. When the surface is allowed to warm, the trapped gas will be released.

Cryogenic pumps are first evacuated to a pressure of 10^{-2} torr, or below, by a mechanical pump. They must begin operation at relatively low pressure, both to preserve the coolant and to prevent the formation of too thick a layer of condensed gas on the surface. Thus, the cryogenic pump should not be operated at pressures greater than 10^{-2} torr.

Cryogenic pumps are capable of reaching very low pressures: Systems capable of reaching 10^{-14} torr have been developed. The pumping speed can be high, with commercially-available units offering pumping speeds up to 5000 liters/second.

VACUUM SYSTEMS

The pump is only one component, although an important one, in a complete vacuum system. In addition to the pump, the system will contain components such as tubes, baffles, cold traps, valves, and vacuum gages. (The gages used to measure vacuum levels are described in Module 10F5.)

The materials from which the vacuum system is constructed are important. They must be materials having low vapor pressure. The lowest pressure that a vacuum system can reach is equal to the vapor pressure of the materials from which it is made. The materials also must possess mechanical strength sufficient to withstand the pressure difference between the inside and the outside of the system. For adequate strength, the walls of the vacuum chamber must be reasonably thick. Also, the materials should have the capacity to be easily cleaned.

To satisfy these requirements, vacuum systems are constructed either of glass or of metal. Stainless steel often is utilized because of its resistance to corrosion; but when cost is an important consideration, aluminum or brass may be used. For ultrahigh vacuum work, stainless steel almost always is the choice.

For lubrication of moving parts, some vacuum greases that have vapor pressures as low as 10^{-8} torr are available. For ultrahigh vacuum work, all greases and waxes are to be avoided.

The materials of the system must be cleaned carefully to avoid release of gases trapped on the surface. After assembly, the system often will be baked while pumping at temperatures ranging up to 400°C, to remove entrapped gases.

In systems using diffusion pumps, the backstreaming of oil molecules into the vacuum chamber is undesirable. Baffles and cold traps may be used to reduce this flow. Baffles are designed to ensure that any molecule moving back toward the chamber must strike at least once on a surface. The surfaces may be water-cooled to aid condensation of the oil on the surface.

Cold traps are cooled by liquid air. Gas passing through a cold trap must make at least two right-angle turns. Oil vapor striking the cold surfaces will condense.

Valves frequently are used to isolate portions of a vacuum system. This technique often is employed when two pumps are used to evacuate a system. Suppose, for example, that both a mechanical pump and a sputter-ion pump are employed to evacuate a particular chamber. When the chamber is allowed to refill with air, the sputter-ion pump may be sealed off behind a valve and allowed to remain in vacuum. When the system is evacuated again, it is first pumped by the mechanical pump. When the limiting pressure of the mechanical pump is reached, the valve is reopened and the sputter-ion pump further reduces the pressure. This technique allows more efficient use, and less contamination of, the sputter-ion pump.

APPLICATIONS OF VACUUM

Vacuum technology has many applications in science and industry. Most scientific laboratories routinely use vacuum equipment to conduct research in areas such as electrical conduction in low-pressure gases, plasma physics, electron microscopy, and ultraviolet spectroscopy.

In more routine applications, vacuum equipment is used in industrial production for many purposes. Light bulbs first are evacuated, then refilled with an inert gas to increase their lifetime. Electronic tubes must be evacuated to provide reasonably long paths for the electrons traversing the tube. The electron tubes used in the control and amplifier circuits of TV sets and picture tubes alike depend upon vacuum technology.

Since vacuum provides good resistance to heat flow, double-walled flasks, with the space between the walls evacuated, possess good thermal insulation. This property is used in thermos bottles to keep food warm or cold.

Many foods are vacuum processed to remove humidity: Products such as freeze-dried coffee, fruit juice, and condensed milk are produced with the aid of vacuum technology.

Metals evaporated within a vacuum system produce reflective coatings on many different items. Plastic toys and jewelry are coated with a thin film of aluminum for decorative purposes. The aluminum is evaporated from tungsten filaments heated electrically. When the vaporized metal strikes a surface, it deposits on the surface and gradually builds up a reflective layer. Such operations generally are carried out at a pressure of about 10^{-3} or 10^{-4} torr. Vacuum deposition of metals probably is the largest single use of vacuum technology in industry.

The range of applications described above represent only a few of the areas in which vacuum technology is important.

SUMMARY

A number of different pumps are used as energy converters to produce a vacuum in enclosed chambers. These pumps include rotary mechanical pumps, diffusion pumps, sputter-ion pumps, and cryogenic pumps. The pressure ranges that can be covered vary from one pump type to another. In common practice, a mechanical pump is used to pump down from atmospheric pressure to medium vacuum; and then a second pump is switched on to reach high or ultrahigh vacuum.

A pump for reducing pressure in a chamber may be characterized by its pumping speed. The pumping speed measures the rate at which gas is removed by the pump. It has units of volume per unit time; for example, liter/second. The definition of pump speed is the volume of gas per unit of time removed from the system at the pressure existing at the inlet to the pump.

VACUUM PUMP EVACUATION OF VACUUM CHAMBER

MATERIALS

Vacuum system, including:
- Mechanical pump.
- Either a diffusion pump or a sputter-ion pump.
- Vacuum gages; for example, a mechanical gage plus a thermocouple gage and an ionization gage (or other equivalent combinations).
- Valves between the pumps and the chamber to be evacuated.
- Flanges, pipes, coupling, as required.

A clock or watch with which to measure time.
Semilogarithmic graph paper.
Stopwatch or lab timer.

PROCEDURE

The student will learn to use vacuum pumps and will use at least two different pumps to evacuate a vacuum system. The vacuum system will be assumed to be available. The student is not requested to set up a vacuum system since that is beyond the scope of this module.

The pressure will be monitored during pump down by vacuum gages, the operation of which is assumed to be known from Module 10F5. Familiarity with the use of a mechanical gage, of a thermocouple gage, and of an ionization gage is assumed. Other gages may be substituted.

A sketch of a system in presented in Figure 6. The first pump is a mechanical one. The second pump may be either a diffusion pump or a sputter-ion pump. In systems where the mechanical pump pumps through a diffusion pump, the procedure will be varied accordingly. (See instructor.)

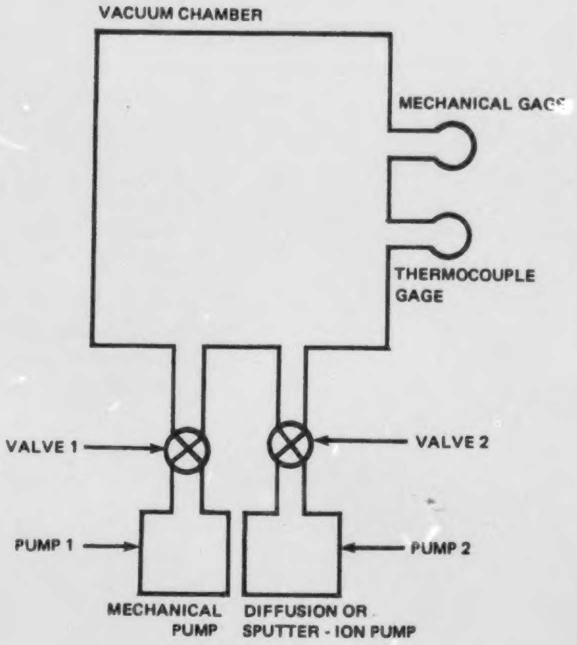


FIG. 6

1. Open valve 1, and close valve 2.
2. Turn on mechanical pump.
3. Observe pressure on the mechanical gage, and record in Data Table the pressure at various times after pump is started.
4. When pressure range appropriate to thermocouple is reached, turn on the thermocouple gage. (See Module 10F5.) Record pressure versus time.
5. When pressure has stopped decreasing, record ultimate pressure reached by mechanical pump.
6. CAUTION: Some precautions are necessary to prevent overheating of a sputter-ion pump if one is being used. Valve 2 must be opened very slowly. If sputter-ion pump begins to overheat, close valve again, and allow it to cool. Now close valve 1, and open valve 2. Start pump 2.
7. With the ionization gage, measure pressure in system at various times after pump is turned on.
8. Record ultimate pressure reached by system.
9. Plot pressure versus time from Steps 7 and 8. The plot may be made on semilogarithmic graph paper.
10. From plot, choose two values of pressure, and measure time required for the pump to reduce the pressure from the higher value to the lower value.
11. Measure and record dimensions of chamber and of pipe.
12. From Steps 10 and 11, calculate pumping speed of the system (pump plus pipe). Use Equation 1.
13. Calculate conductance of the pipe (Equation 6), and then calculate pumping speed of pump (Equation 7). Compare result with manufacturer's specified pumping speed. Is there a difference between the result and the specified speed? If so, why?

DATA TABLE

TYPE OF PUMP

Mechanical			Diffusion or Sputter-ion								
Time Interval	Pressure	Ultimate Pressure	Time Interval	Pressure	Ultimate Pressure	Time required to pump from 1st pressure to 2nd (from plot)	Dimensions		Pumping Speed of System	Conductance of Pipe	Pumping Speed of Pump
							Vacuum Chamber	Pipe			
						The two plotted pressures: 1. ____________ 2. ____________					

SELF CHECK

1. Define the following terms:

   Pumping speed.
   Pump-down time.
   Conductance for flow of gas through a pipe.
   Impedance for flow of gas through a pipe.

2. A pump has a pumping speed of 100 liter/second. How much time will be required for that pump to reduce the pressure in a 100-liter chamber from 1 torr to 10^{-5} torr?

3. A pipe has a diameter of 5 centimeters and a length of 50 centimeters. What is the impedance of the pipe, when the pressure is 10^{-3} torr?

4. Describe the operation of the following pumps:

   Rotary mechanical pumps.
   Diffusion pumps.
   Sputter-ion pumps.
   Cryogenic pumps.

5. State the approximate pressure range covered by each of the above pump types.

6. What is the total impedance of two pipes with individual impedances of 0.5 sec/liter and 0.1 sec/liter if the pipes are connected in parallel?

7. What is the diameter of a pipe that is 40 cm long and has a total conductance of 10 liters/second if the pipe is circular and Equation 6 applies?

8. What is the conductance of a pipe if the total pumping speed through the pipe is 10 l/sec and the pumping speed of the pump alone is 20 l/sec?

9E1

ELECTRIC GENERATORS

The conversion of mechanical energy into electrical energy is perhaps the most significant form of energy conversion. Today's technological society depends greatly upon the ability to generate, transport, store and utilize electrical energy. An electric generator is an energy converter that changes rotational mechanical energy to electrical energy. This module presents the theory of operation and basic design of d.c. generators. In the laboratory, the student will operate a d.c. generator and will measure the electrical power it produces.

OBJECTIVES

Upon completion of this module, the student should be able to:

- Use the right-hand generator rule to determine the direction of force on electrons in a conductor moving through a magnetic field.
- Determine the direction of electron flow, given a diagram of a loop of wire rotating in a magnetic field.
- List the three factors that determine the magnitude of the voltage produced by a genrator.
- Draw and label sketches showing the differences in d.c. and a.c. generators and the wave forms produced by each.
- State two methods commonly used to produce the magnetic field in generators.

- Given the appropriate equipment, operate a d.c. generator, and measure the electrical power produced.

DISCUSSION

The magnetic force acting on an electric charge moving through a magnetic field was discussed in Application Module 1E4, "Magnetic Forces and Fields." This module describes the application of these forces in generation of electric currents.

Figure 1 illustrates the right-hand generator rule used to predict the direction of force on an electron in a wire. The forefinger points in the direction of the magnetic field. The middle finger points in the direction of motion of the conductor. The thumb points in the direction of the force (or thrust) on the electron and, thus, the direction of electron flow in the wire.

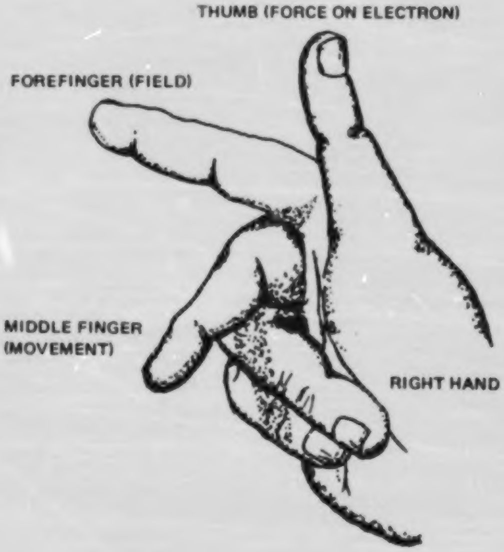
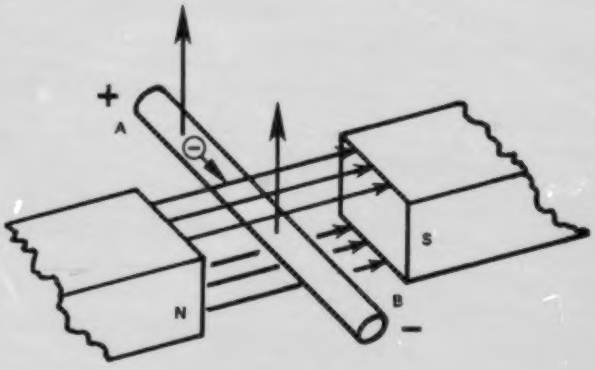


FIG. 1

Figure 2 shows the application of this rule to a straight piece of wire moving through a uniform field between two magnetic poles. In Figure 2a the field is to the right and the wire moves upward. Magnetic force moves electrons in the wire from A toward B, resulting in a potential difference across the wire as shown. In Figure 2b the conductor moves downward, reversing the direction of force on the electron and the polarity of the voltage across the wire.

In Figure 2c the conductor moves along lines of magnetic force, in which case charged particles in the wire do not cross magnetic-field lines; and there is no force on the electron. Thus, no voltage is present across the conductor.

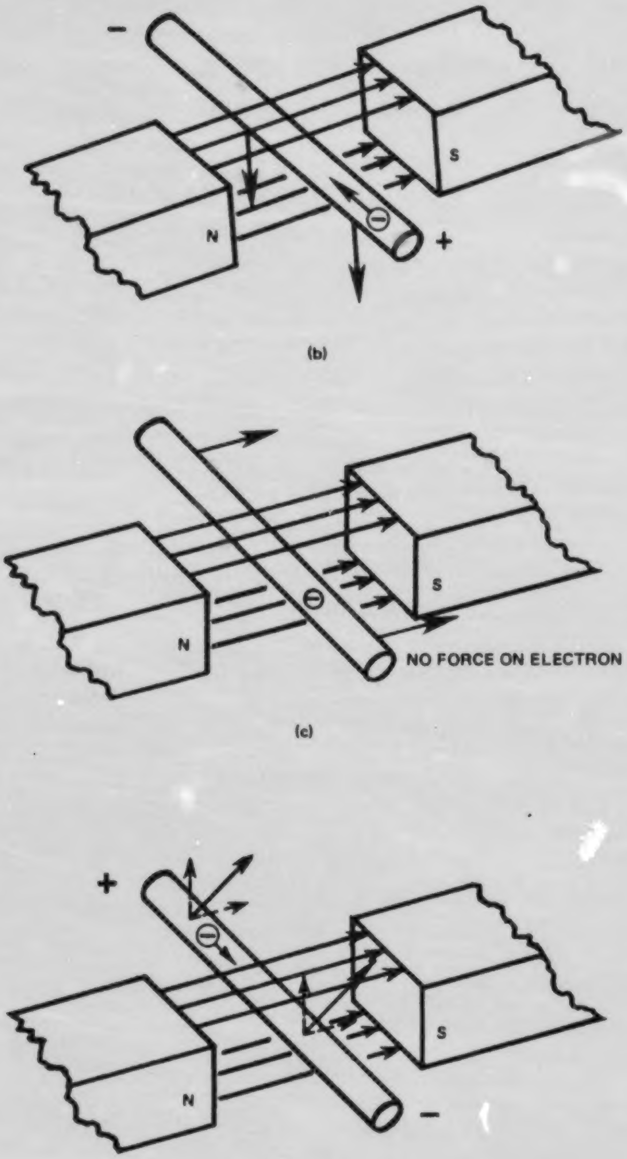
Figure 2d shows the results when the conductor moves at an angle other than perpendicular or parallel to the field lines. In this case, the vertical component of the motion produces a force on the electron just as the vertical motion in Figure 2a. The horizontal component produces no force on the electron.



(a)

FIG. 2

Figure 3 shows how this effect is applied in a d.c. electric generator. This generator consists of a square

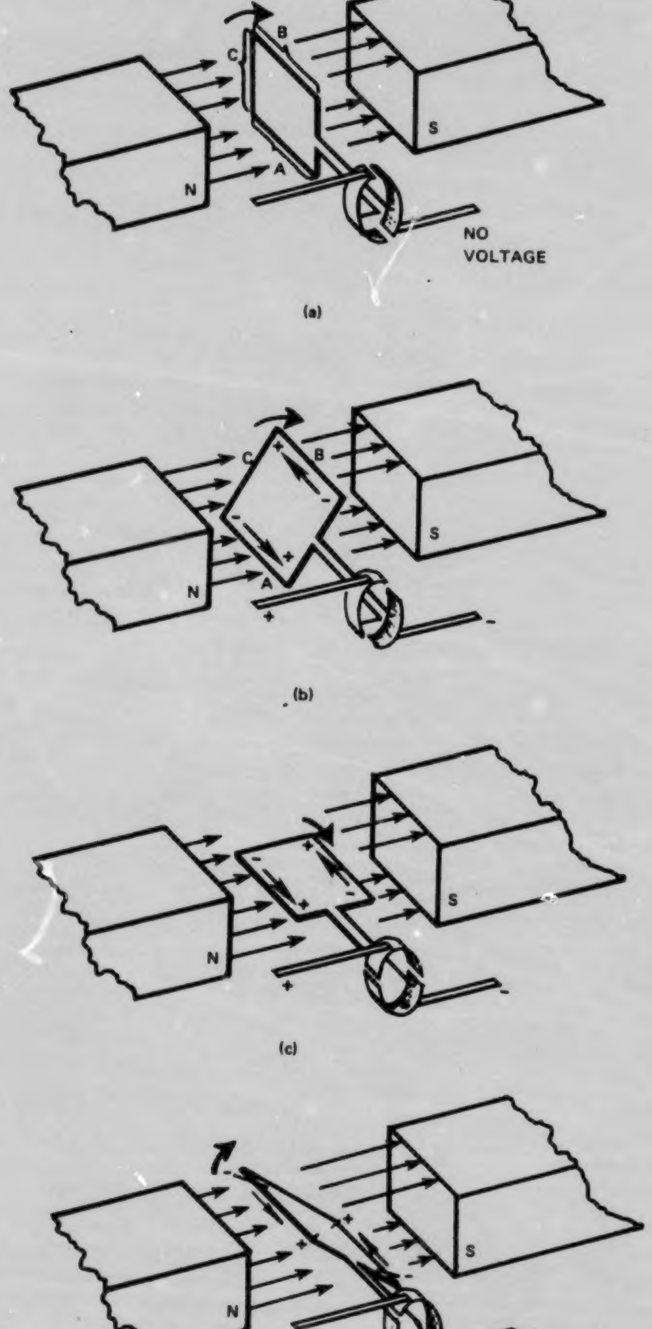

–
S
N
+
NO FORCE ON ELECTRON
S
N
+
S
N
–

(b)

(c)

(d)

FIG. 2


C
B
A
N
S
NO VOLTAGE
+
-

(a)

(b)

(c)

(d)

FIG. 3

loop of wire rotating in a fixed magnetic field and of a split ring commutator.

In Figure 3a, the upper output contact (called a brush) is connected to wire segment A. At this time, both wire segments are moving parallel to the magnetic-field lines, and no voltage is generated. This posiiton is point a on the graph in Figure 4. This graph shows the voltage across the generator brushes as a function of time.

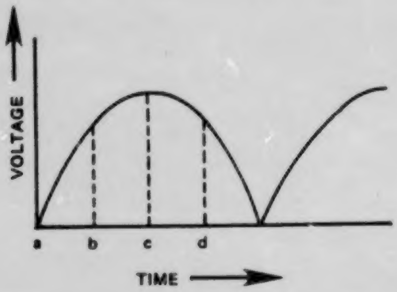


FIG. 4

Figure 3b shows the position of the loop some short time later. Wire segments A and B are now moving at an angle to the magnetic-field lines, and a voltage is being produced in sides A and B. Side B is moving downward through the field, causing a force on the electrons in the direction indicated by the arrow. Side A moves upward, resulting in a force in the opposite direction. These two voltages add in series to produce the output voltage shown as point b in Figure 4. Wire segment c contributes no voltage, because it crosses no magnetic field lines as it moves.

In Figure 3c, the two voltage-producing wire segments are moving perpendicularly to the direction of the magnetic field. This position produces the maximum voltage, indicated by point c in Figure 4.

In Figure 3d, the motion of the wire segments is no longer perpendicular to magnetic field lines, and the output voltage has dropped to that indicated by point d in Figure 4. When the loop reaches the vertical position again, the voltage will be zero. Each brush will slip to the other segment of the commutator ring, and the entire process will be repeated.

The amount of voltage present across the wire segment depends upon the strength of the magnetic field and the velocity of the segment perpendicular to the field. This voltage is small for practical fields and velocities. Generation of usable voltage requires a rotating coil of many loops, shown schematically in Figure 5. As this coil rotates, a voltage is produced across each hroizontal wire segment. All these voltages add in series to produce the total output voltage of the generator. This output is shown as a function of time in Figure 6. The amplitude (height or maximum voltage) of this curve is determined by three factors:

- Number of loops in the moving coil.
- Strength of the magnetic field.
- Rotational rate of the coil.

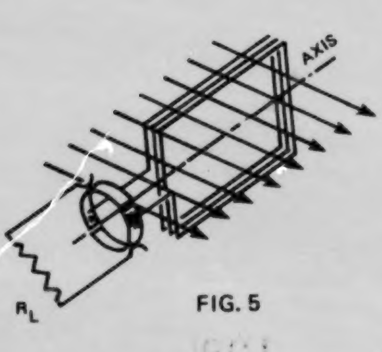


FIG. 5

When any one of these factors is increased, output voltage also increases. If the rotational rate is increased, the frequency at which the output varies is also increased.

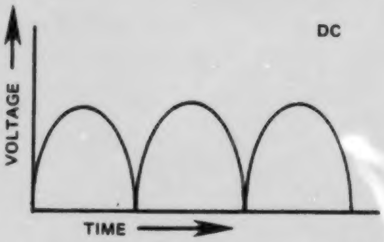


FIG. 6

If the same generator is equipped with a pair of slip-ring contacts as shown in Figure 7, an a.c. voltage will be produced. The output of such a generator, often called an "alternator" because it generates an alternating current, is shown in Figure 8.

Permanent magnets often are used to provide the magnetic field in small generators, but larger units require fields of greater strength than those available from permanent magnets. Figure 9 shows how this problem is solved in most generators. A d.c. current flows through a stationary field coil to produce a magnetic field. A multipole armature, shown in Figure 10, is rotated in this magnetic field to produce the output voltage. This armature consists of several coils of wire, each connected to two opposing segments of the commutator. This arrangement allows only the maximum voltage portion of the voltage-versus-time curve of each coil to be connected to the generator output. The resulting output voltage is shown in Figure 11.

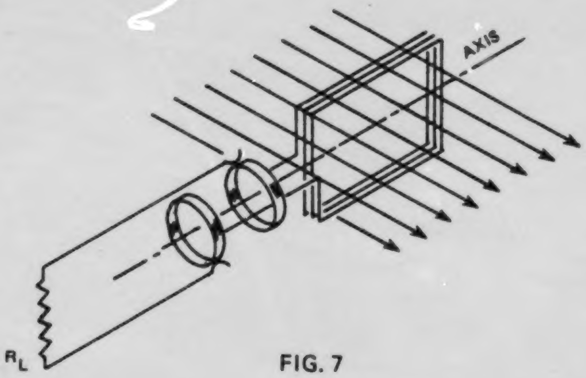


FIG. 7

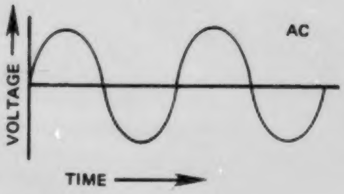


FIG. 8

When a generator is turning with a load connected to the output, the current flowing in the load circuit is also flowing in the armature. Such current establishes a magnetic field around the armature; this field opposes the fixed magnetic field of the generator. The interaction of these two magnetic fields, one fixed and one rotating, opposes the rotation of the armature; therefore, rotational mechanical energy must be supplied to the armature to maintain its rotation. Thus, the armature of a generator rotates easily if no load is attached, but requires greater torque as the current delivered increases. This situation is required by the principle of the conservation energy. The maximum current that may be produced by a generator is limited by the size of wire used in the armature.

The most important application of d.c. generators arises from the fact that d.c. electric motors allow finer speed control than either a.c. motors or engines that derive their power from the combustion of chemical fuels. In locomotives and ships, diesel engines or steam turbines drive d.c. generators; electrical output power then is used to operate d.c. motors that provide propulsion.

SUMMARY

Electric generators (and alternators) convert rotational mechanical energy into electrical energy. Magnetic forces and fields make the operation of generators possible.

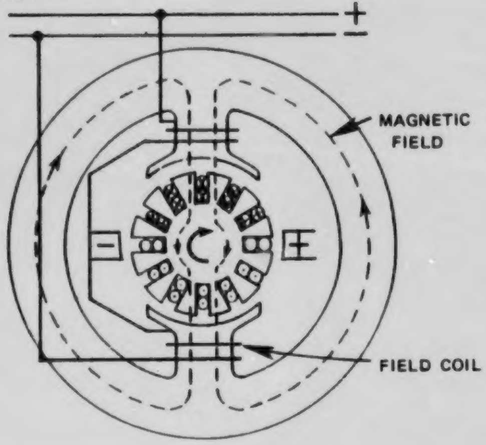


FIG. 9

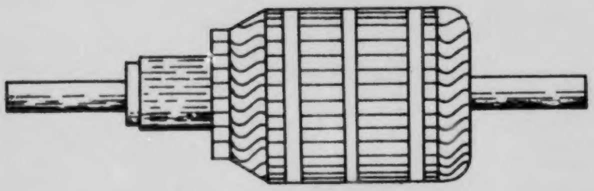


FIG. 10

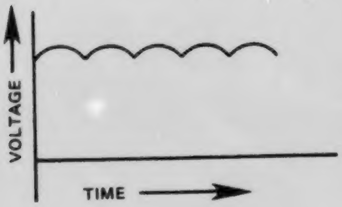


FIG. 11

GENERATOR OPERATION AND CHARACTERISTICS

MATERIALS

Demonstration motor-generator (Sargent-Welch #2469 or equivalent).
Two variable d.c. power supplies, 0-6V, 5A with meters to indicate output current and voltage.
d.c. voltmeter, 0-25V.
d.c. milliammeter, 0-500 mA.
Single-pole, single-throw switch.
Wire for connectors.
10 Ω, 1W resistor.

PROCEDURE

1. Set up equipment as shown in Figure 12.

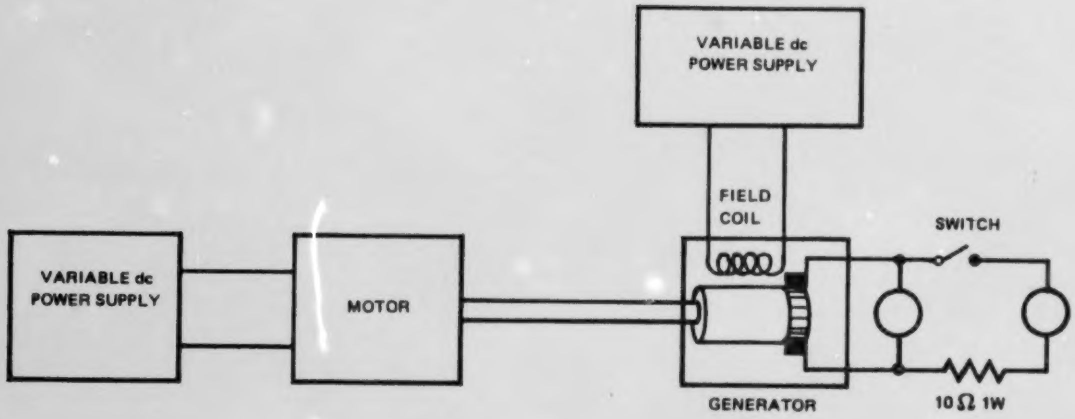

VARIABLE dc POWER SUPPLY
FIELD COIL
SWITCH
VARIABLE dc POWER SUPPLY
MOTOR
GENERATOR
10 Ω 1W

FIG. 12

2. Turn on d.c. power supply connected to motor, and set voltage to produce a moderate motor speed.

3. With switch open, turn on d.c. power supply connected to generator field coil, and set voltage at approximately 3V.

4. Measure output voltage of generator. Record field-coil voltage, output voltage of generator, motor voltage and motor current in Data Table.

5. Increase field-coil voltage to 6V, and repeat step 4.

6. Using input voltage of motor as in step 2, close switch.

7. Repeat entire procedure, recording all data as before with addition of output current of generator.

8. Determine input power of motor and output power of generator from P = I V and complete Data Table.

9. Why is more power required to turn a larger generator (or one with larger windings) than to turn a smaller one?

DATA TABLE

NO LOAD

Motor			Generator	
Voltage (V)	Current (A)	Power (W)	Field coil voltage (V)	Output voltage (V)

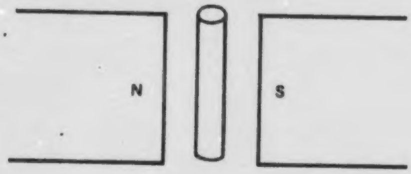
LOAD

Motor			Generator			
Voltage (V)	Current (A)	Power (W)	Field coil voltage (V)	Output voltage (V)	Output current (MA)	Output power (W)

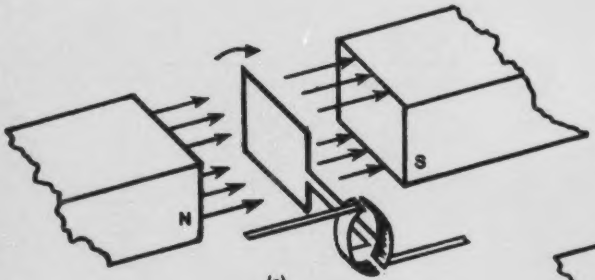
SELF CHECK

1. Determine the direction of force on electrons in the wire at the right if the wire moves

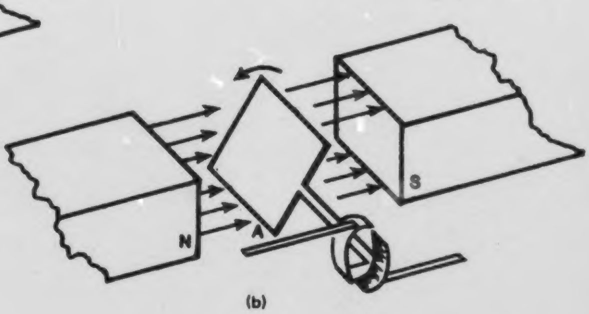
   a. Out of the page toward the reader.
   b. Toward the south pole of the magnet.



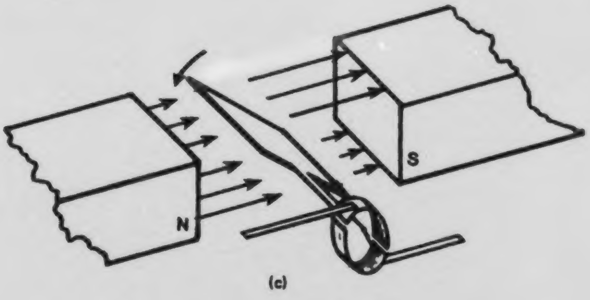
2. Indicate the direction of electron flow through the generator coil in the following diagrams. (Use the right-hand generator rule for electron flow presented in the text.)



(a)



(b)



(c)

3. List the three factors that determine the voltage produced by a generator.

4. Sketch the output waveforms from an a.c. generator and a d.c. generator.

5. State two methods of producing magnetic fields in generators.

6. Calculate the input power to a motor if the input voltage and current are 110 V and 1.6 A respectively.

7. Calculate the power output of the generator of problem 6 if the output voltage and current are 110 V and 1 Amp. What is the efficiency of the generator?

8. What is the effective resistance of a device powered by a small electric generator which has an input power of 400 W and is 50% efficient if a current of 2 Amps flows through the device?

9. How much charge flows out of a generator in 2 hours if the power put into the generator is 3000 watts, the generator has an efficiency of 47% and the output voltage is 110 volts?

10. Four identical electric generators are connected in series. (That is the output of the first generator is the input of the second, and so on.) If 4000 W of power are input to the first generator and 500 watts of power come out of the fourth generator, what is the efficiency of each of the generators? What assumption concerning input energy and efficiency must be made in order to solve this problem?

9E2

ALTERNATORS

Generator principles discussed in Module 9E1, "Electric Generators," apply to alternators as well as to d.c. generators; both convert mechanical energy into electrical energy. Alternators produce an a.c. waveform and are used in power-generating plants to provide large quantities of electrical energy for home and industrial uses. Rectifying diodes are used to produce d.c. voltage with an alternator. Automobile alternators convert rotational mechanical power to d.c. voltage and current by this means. This module presents the theory, construction details, and applications of alternators. In the laboratory, the student will determine the output power and efficiency of an alternator.

OBJECTIVES

Upon completion of this module, the student should be able to:

- Explain with the use of diagrams how an alternator produces an a.c. voltage.

- Draw and label diagrams showing the construction of the rotor of an alternator having several pole pairs. Include the electrical windings, rotor shell, brushes, and slip-rings.

- Given any two of the following quantities, calculate the third:

  Number of pole pairs.
  Rotational rate in revolutions per second.
  Output frequency.

- State two major advantages of alternators over generators for the production of electrical power.

- Given the number of pole pairs in the rotor of a three-phase alternator, determine the total number of coils in the stator; and explain how they are connected.
- Given the appropriate equipment, measure the efficiency of an automobile alternator; and observe the output waveform.

DISCUSSION

A review of the electrical generator principle (Module 9E1) will help in the understanding of alternator operation. Recall that to generate electrical energy from mechanical energy, relative motion must exist between a magnetic field and a conductor. When alternators are used to generate electrical energy, the magnet moves in a rotary motion about a fixed coil. All the rules discussed in Module 9E1 for electrical generators also apply to alternators.

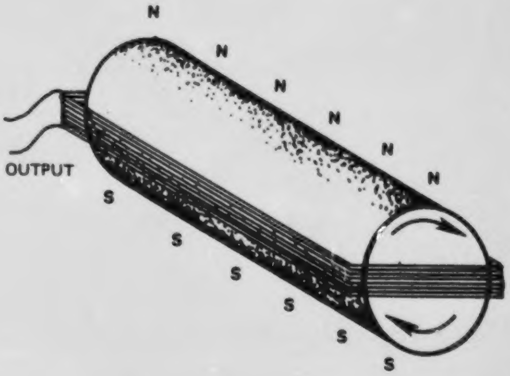


FIG. 1

Figure 1 shows a magnet that rotates inside a coil to produce voltage. The direction of induced voltage at successive points of rotation is shown in Figure 2. The direction of induced voltage is the same as the direction of conventional current flow. Note that in Figure 2a the rotating magnetic field is not cutting the coil, and that no voltage is induced, therefore, into the coil at this point. During the first quarter turn, the voltage increases and reaches a maximum as the poles of the magnet pass the coil of wire and decreases during the next quarter turn. In Figure 2d, the magnet has rotated a full 180°, or one-half turn, and the induced voltage is again zero. The direction of current flow in the coil can be determined by the left-hand generator rule and is shown by

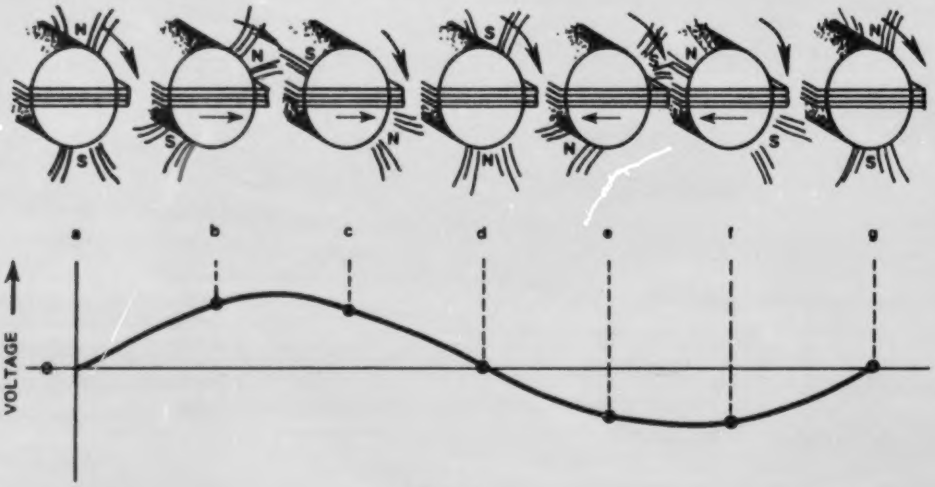


FIG. 2

the arrows.

One complete rotation of the magnet produces one cycle of induced alternating voltage, as shown. When the magnet is rotated at a rate of 3600 revolutions per minute, or 60 revolutions per second, the frequency of the waveform produced is 60 cycles per second, or 60 Hz. The angular velocity of the rotating magnet in industrial alternators is very closely controlled, to maintain the exact frequency of 60 Hz, which is the standard frequency of the a.c. voltage used in homes and industry.

Magnets having four or more poles generally are used in industrial alternators. Since there are more pairs of poles producing magnetic fields cutting the coils of wire, more complete waveforms of a.c. voltage are produced for each rotation of the rotor. The relationship between number of magnetic pole pairs, frequency, and angular velocity of the rotating magnet is expressed by Equation 1.

$$f = N\omega \tag{1}$$

where

f = frequency of induced voltage in Hz.
N = number of pole pairs.
ω = angular velocity of the rotating magnet in revolutions per second.

Examples A and B show how Equation 1 is used in solving problems.

EXAMPLE A: ALTERNATOR OUTPUT FREQUENCY

Given: An alternator has four pole pairs and an angular velocity of 15 revolutions per second.

Find: Frequency of the waveform produced.

Solution:

$$f = N\omega$$
$$= (4 \text{ pole pairs})(15 \text{ revolutions per second})$$
$$f = 60 \text{ Hz}$$

EXAMPLE B: DETERMINATION OF ALTERNATOR ANGULAR VELOCITY

Given: An alternator has 10 pole pairs and frequency of 60 Hz.

Find: Revolutions per second (rps).

Solution:

$$f = N\omega$$
$$\omega = \frac{f}{N}$$

$$\omega = \frac{60\ \text{Hz}}{10\ \text{pole pairs}}$$

$$\omega = 6\ \text{rps}$$

The relationship between frequency, pole pairs, and alternator angular velocity shown in Examples A and B is used in actual alternator applications. Most steam turbines operate efficiently at high angular velocity, allowing single pairs of magnetic poles to be employed in alternators driven by steam turbines. Water-driven alternators rotate at a lower velocity and, therefore, contain a greater number of magnetic pole pairs. Figure 3 shows a rotating magnet (rotor) having several pairs of magnetic poles. Each pole is produced by a magnet.

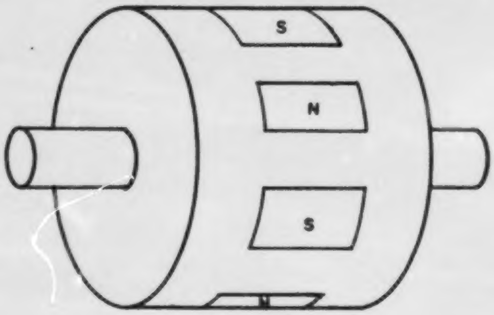


FIG. 3

Permanent magnets can be used in alternators to establish the magnetic lines of force, but this is seldom done in actual practice. Recall from Module 9E1 that the magnitude of voltage produced by a generator can be regulated by varying the strength of the magnetic field that cuts the conductor. An electromagnet must be used in order that strength of the magnetic field can be changed by current adjustment in the electromagnet. When the magnetic coil illustrated in Figure 4 is placed inside an iron shell, the multiple-pole pairs shown in Figure 5 is the result. The

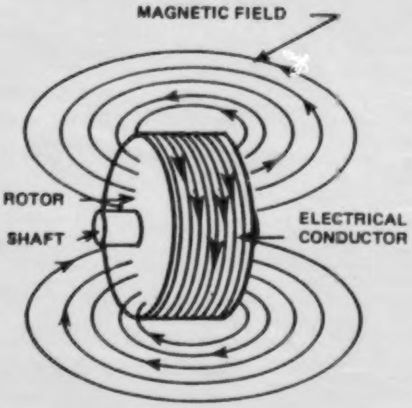


FIG. 4

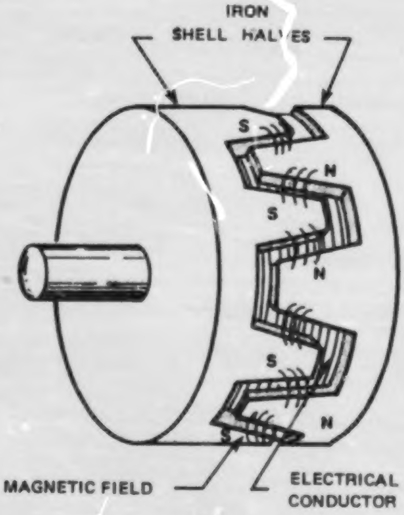


FIG. 5

strength of the many magnetic fields established by the fingers of the iron shell depends upon the amount of current in the inner coil.

At least a small amount of voltage can be generated with a field current of zero. This generation is possible because of the residual magnetism of the armature and may be the case for some automobile alternators. Large industrial alternator applications do not rely upon residual magnetism of the rotor to produce a magnetic field for voltage generation when the alternator is started with no voltage present. Diesel- or oil-driven d.c. generators produce a voltage, called "excitation voltage," to supply the field current for the rotor when the alternator is to be started.

Current to regulate the field strength of the rotating magnet flows through the slip-rings and brushes, as shown in Figure 6. This current, known as the "field current," is much less than the current in the "stator," which carries the load on the alternator. Thus, the slip-rings and brushes are less likely to fail with this arrangement than if required to carry the load current, as is the case with the stationary-field and rotating armatures of the d.c. generator. The component parts of an automobile alternator are displayed in Figure 7.

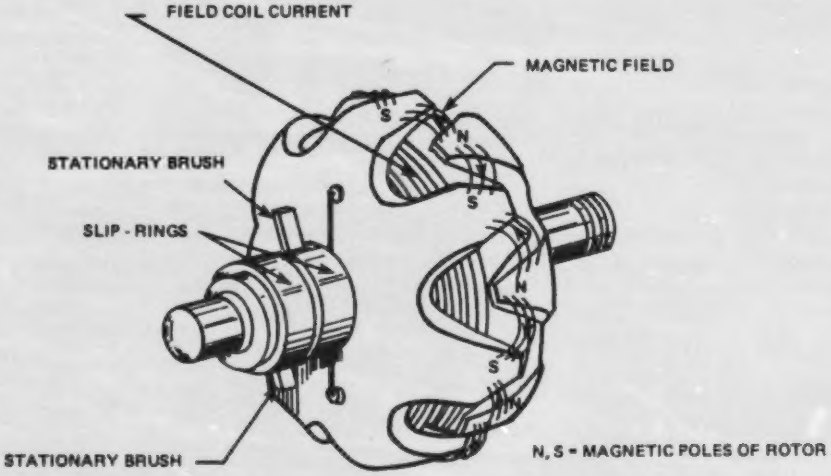


FIG. 6

Automobile alternators contain as many as seven or more pole pairs. In order for the multiple pole arrangement to function effectively, each pair of poles must have a coil in the stator

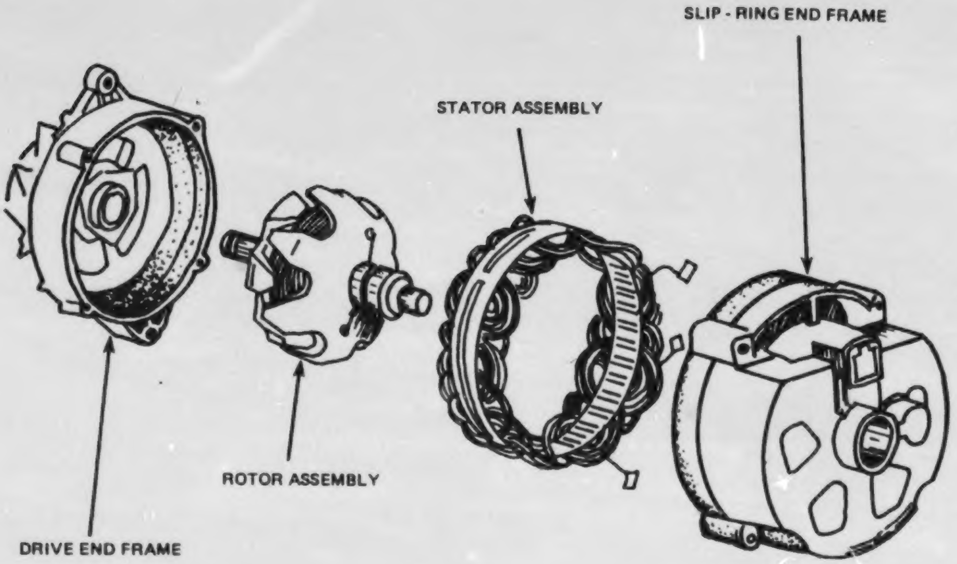


FIG. 7

assembly. All the coils are connected in series to produce an output voltage equal to the product of the number of coils times the voltage produced by a single coil; that is, if an alternator contains seven pole pairs, the voltage output will be seven times as great as with 1 pole pair. Alternators are used on later-model automobiles to generate larger amounts of electrical power than the d.c. generators that were used previously.

In order to increase the electrical power produced by an alternator, coils in the stator are arranged in such a manner that the generation of a waveform is begun every 120° of rotor rotation. The resulting waveform is illustrated in Figure 8. When the value of one waveform decreases, another waveform is generated. This process decreases the net voltage variations and makes rectification easier. Such multiphase generation is utilized for

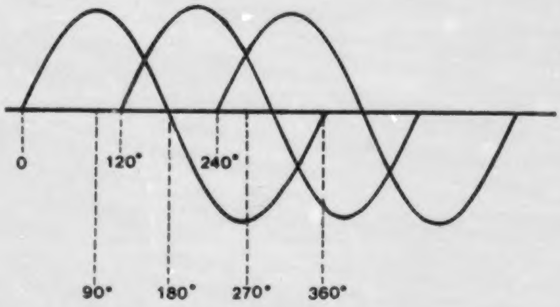


FIG. 8

large electrical devices in industry, to provide more efficient operation.

One set of coils for a seven-pole magnet is displayed in Figure 9. This set of coils produces one of the voltage waves of Figure 8. Two more sets of seven coils are shown in Figure 10, for a total of three sets of seven coils for multiphase generation.

Figure 11 consists of a circuit to rectify the waveform of Figure 8 for d.c. applications. The diodes used for rectification act as electrical check values to allow current to flow in only one direction. Each coil set from the multiphase alternator is connected at points A, B, and C, respectively, of the rectifier circuit. Point D is the common connection of the opposite ends of the three coil sets to the negative output terminal of the alternator. Figure 12 shows the three electrical waves produced before rectification, the result of rectification, and the alternator output voltage as a function of time.

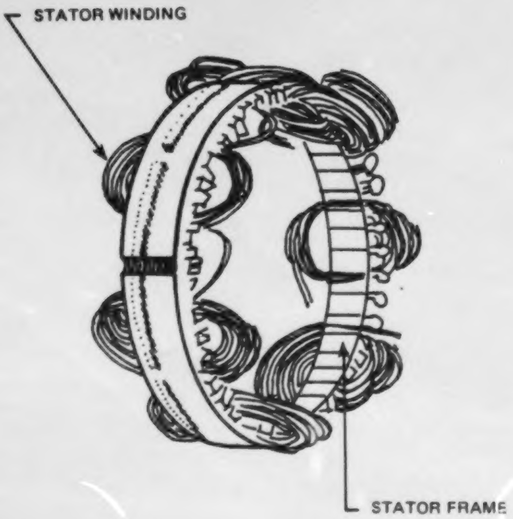


FIG. 9

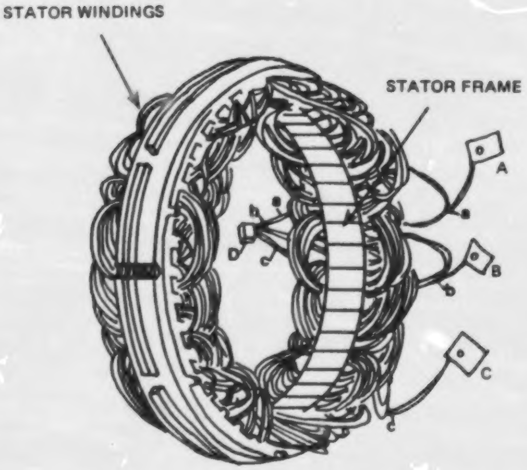


FIG. 10

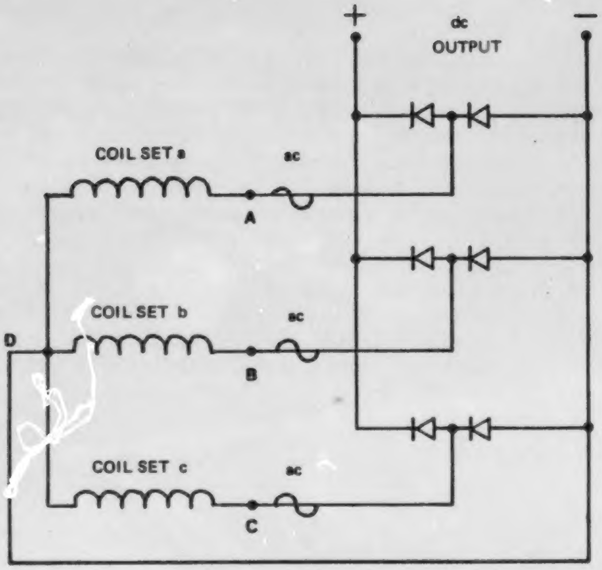

+
dc
OUTPUT
−
COIL SET a
ac
A
COIL SET b
ac
D
B
COIL SET c
ac
C

FIG. 11

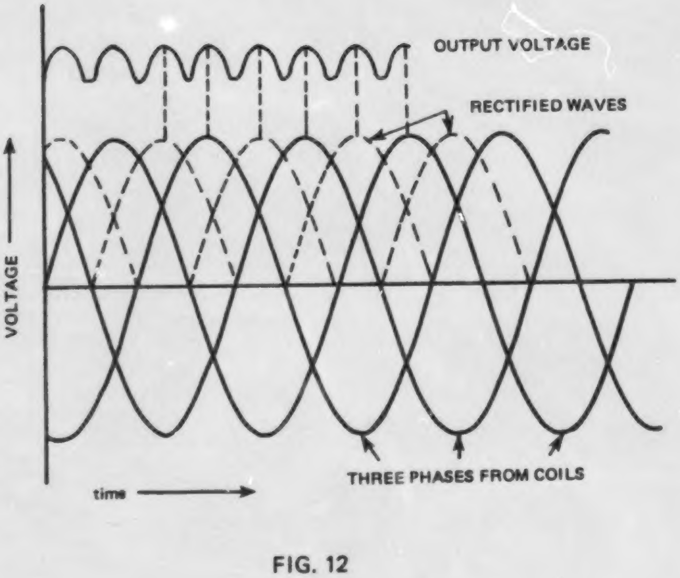

OUTPUT VOLTAGE
RECTIFIED WAVES
VOLTAGE
THREE PHASES FROM COILS
time

FIG. 12

SUMMARY

In an electric generator, a voltage is produced in a conductor that moves through a magnetic field. In an alternator, the voltage is produced as the magnetic field moves through stationary coils. The magnetic field of the alternator rotor is created by a field current carried to a coil through brushes and slip-rings. The much larger output current is carried by the stationary coils that require no brushes. Thus, alternator brushes can be much smaller than generator brushes, and are subject to less wear. Alternator rotors can be much smaller and lighter in weight than generator rotors, because the lower rotor current allows smaller conductors. For maximum efficiency, alternators often are constructed with several magnetic pole pairs in the rotor and three sets of stator windings for three-phase operation. Alternators produce most of the world's electrical power.

OUTPUT POWER OF AN ALTERNATOR

MATERIALS

An alternator (similar to automobile type).
Small variable-speed a.c. or d.c. electric motor to drive alternator.
Variable power supply for motor.
Oscilloscope.
Assorted resistors.
VOM.
3 SPST switches.

PROCEDURE

1. Connect the motor and alternator as shown in Figure 13.

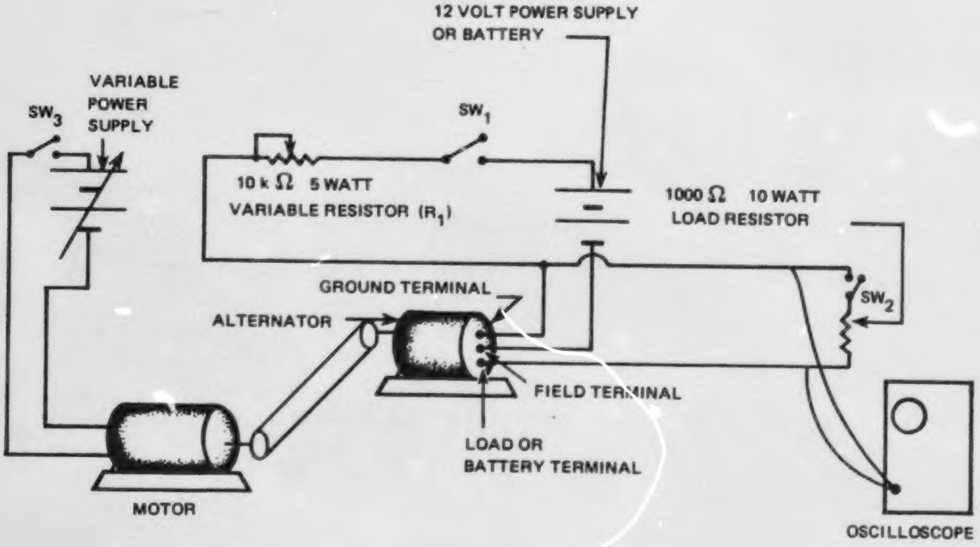

12 VOLT POWER SUPPLY OR BATTERY
VARIABLE POWER SUPPLY
SW_3
SW_1
10 k Ω 5 WATT VARIABLE RESISTOR (R_1)
1000 Ω 10 WATT LOAD RESISTOR
SW_2
GROUND TERMINAL
ALTERNATOR
FIELD TERMINAL
LOAD OR BATTERY TERMINAL
MOTOR
OSCILLOSCOPE

FIG. 13

2. Open switches SW_1, SW_2, and SW_3.

3. Adjust voltage of variable power supply to zero, and close SW_3.

4. Slowly increase the motor voltage until alternator output is barely detected on oscilloscope. Observe and note waveform.

5. Increase motor voltage until motor is turning at a moderate speed. Take appropriate readings, and enter in Column 1, Trial 1, of Data Table.

6. Set R_1 for its maximum value, close SW_1, and slowly increase field current, until an increase in output voltage is detected on the oscilloscope. Measure and record appropriate values in Data Table for Trial 2.

7. Increase field current in three separate increments, and enter appropriate data in Column "A" of Data Table.

8. Open SW_1 and SW_3. Close SW_2. This procedure places a load on the alternator. To study operation of the alternator under load conditions, complete the following steps:

9. Close SW_3, and adjust motor current to produce the same alternator output voltage detected in Step 5. Because of the load on the alternator, a field current may have to be established in order to obtain the desired voltage at a moderate rate of motor speed. If this is the case, set R_1 for its maximum value, close SW_1, and increase the field current for the voltage noted in Step 5.

10. Repeat Step 7, and enter data in Column "B" of Data Table.

11. Observe alternator output waveform on oscilloscope.

12. Increase voltage to motor until a change in motor speed is detected.

13. Observe the waveform.

14. The motor power and generator power can be calculated from data entered into Data Table. (Power = VI.)

15. To calculate overall efficiency, divide power out by power in, and multiply by 100:

$$\eta = \frac{P_{out}}{P_{in}} (100)$$

This equation gives the efficiency of the motor-alternator pair in converting electrical energy to mechanical energy and back to electrical energy.

16. Explain two factors that affect the output of the alternator and the effects of each.

DATA TABLE

Trial	Motor Voltage		Motor Current		Motor Power		Field Current		Alternator Output				Overall Efficiency
									Voltage (V)		Current (A)	Power (W)	
	A	B	A	B	A	B	A	B	A	B	B	B	B
1													
2													
3													
4													
5													

A - No load condition
B - Load condition

SELF CHECK

1. Draw and label diagrams showing a magnet rotating inside a coil of wire and showing the waveform generated in the wire.
2. Draw and label the rotor of an automobile alternator, showing the field coil, the brushes, and the magnetic poles.
3. A three-phase alternator rotates at 360 r/min (rpm) and produces a 60-Hz output. How many pole pairs does its rotor contain? How many coils does its stator contain?
4. Why is most electricity produced by alternators instead of generators?
5. A three-phase alternator has a total of 27 coils in its stator and rotates at 420 r/min (rpm). What frequency is produced?

6. If the angular velocity of a rotating magnet is 20 radians per second, how many revolutions does the magnet make per second?

7. What is the overall efficiency of an alternator that requires an input of 700 J in order to output 500 J?

8. For an alternator operating at maximum electrical power, how many waveforms have been generated when the rotor has gone through four complete revolutions? (Hint: See Figure 8 and discussion.)

9E3

ELECTRIC MOTORS

One of the most essential and versatile energy converters utilized by modern technology converts electrical energy into mechanical energy. The device that accomplishes this conversion is called an "electric motor." The advantages offered by electric motors are numerous - they are mechanically simple, quiet, clean to operate, and relatively efficient.

This module introduces the principles governing the operation of direct-current (d.c.) electric motors. In the laboratory, the student will build and operate a permanent-magnet motor and a series motor.

OBJECTIVES

Upon completion of this module, the student should be able to:

- Use the left-hand motor rule to determine the direction of the force on a current-carrying conductor in a magnetic field.
- Given a diagram of a loop of wire carrying a current in a magnetic field, determine the direction of loop rotation.
- Sketch a permanent-magnet d.c. motor. Label and explain the function of the principal parts.
- List the three factors that determine the magnitude of the torque produced by an electric motor.
- State the difference between permanent-magnet, series, and shunt d.c. electric motors.

- Given the appropriate equipment, build and operate a simple permanent magnet motor, and show how it can be converted into a series motor.

DISCUSSION

The rotation of an electric motor is caused by the forces that act upon current-carrying wires in the motor because of the presence of a magnetic field. Operation of an electric motor is made possible by the nature and direction of these forces.

The direction of the force acting on a current-carrying wire in the presence of a magnetic field is determined by the Left-hand Motor Rule. Figure 1 shows how the motor rule is applied. The thumb, forefinger, and middle finger of the left hand are held at right angles to each other. If the forefinger is pointed in the direction of the magnetic field B, and the thumb is pointed in the direction of electron flow, then the middle finger will point in the direction of the magnetic force acting on the wire. Hence, the force on the wire in Figure 2a is straight up. If the direction of electron flow is reversed as in Figure 2b, the force on the wire is in the opposite direction. If the direction of electron flow is parallel to the magnetic field, there is no magnetic force acting on the wire (Figure 2c).

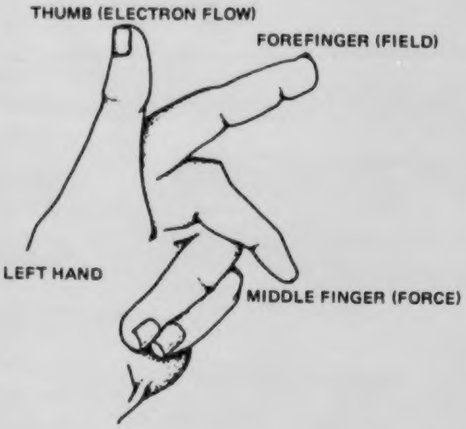
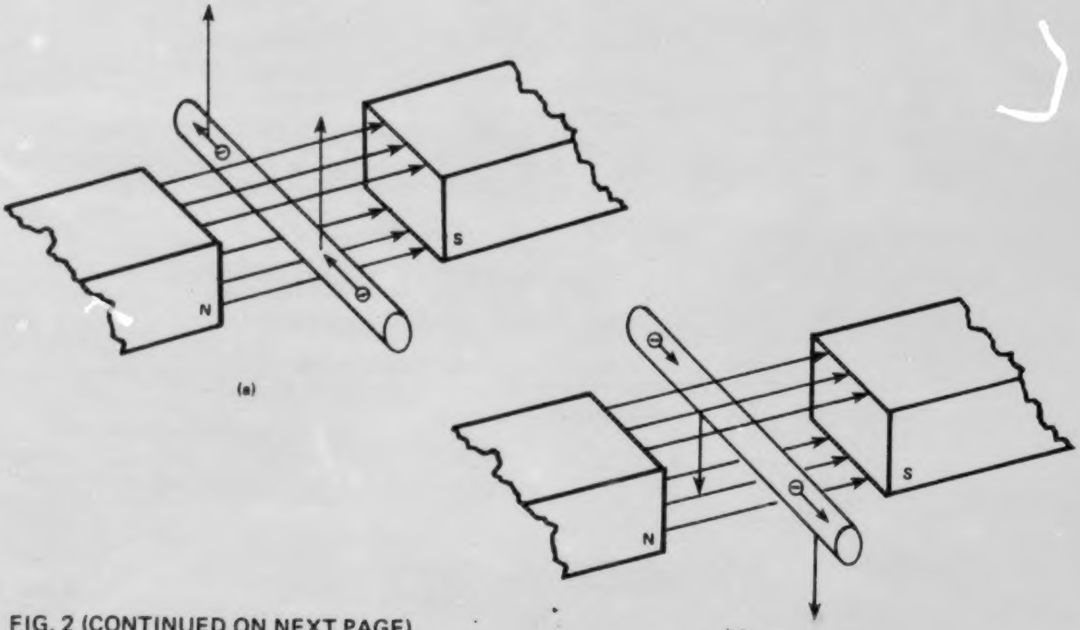


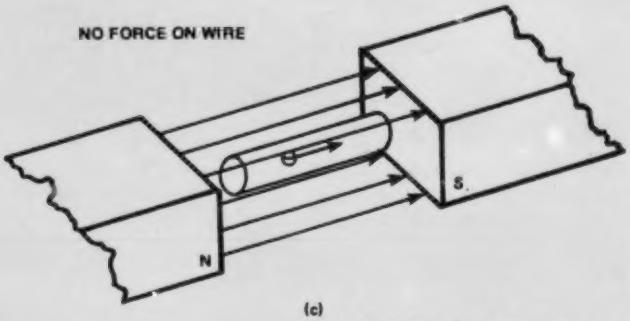
FIG. 1



(a)

(b)

FIG. 2 (CONTINUED ON NEXT PAGE)



(c)

FIG. 2 (CONTINUED FROM PREVIOUS PAGE)

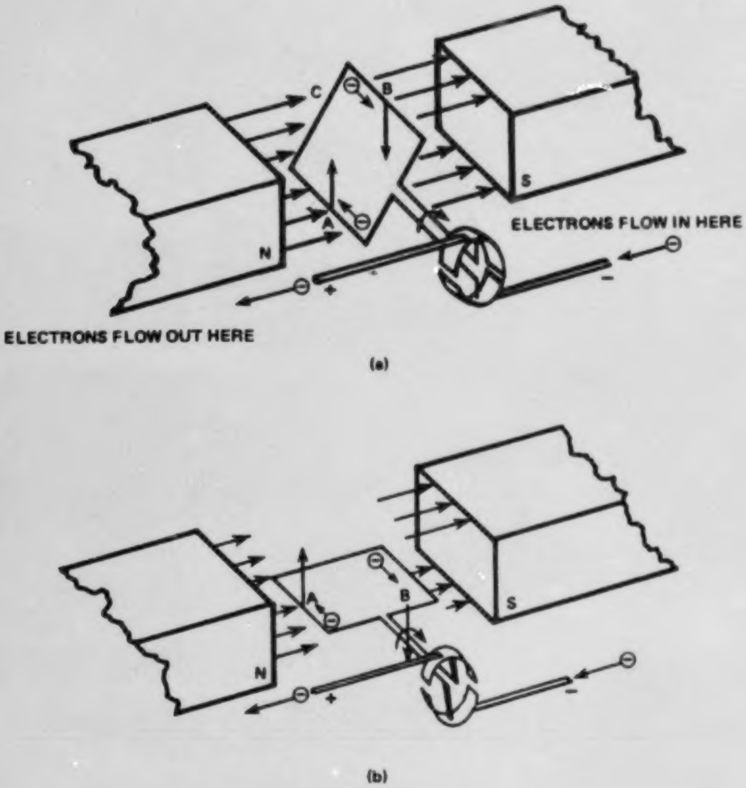


FIG. 3 (CONTINUED ON NEXT PAGE)

Figure 3a shows how a simple electric motor is constructed from a square loop of wire, called the "armature," and from a "split-ring commutator." Because the magnetic field is provided by two permanent magnets, this configuration is called a "permanent-magnet motor."

In Figure 3a, electrons flow into the loop through the upper contact, or brush, and out through the lower brush. Magnetic force is directed upward on Segment A and downward on Segment B. These oppositely-directed forces exert a net torque on the armature, causing it to rotate in a clockwise direction. This torque on the armature is plotted as Point "a" on the graph of Figure 4. Magnetic forces also act on wire Segment C; these forces, however, are always parallel to the axis of rotation and, therefore, never contribute to the torque acting on the armature. For this reason, the forces on Segment C are umimportant and will be disregarded.

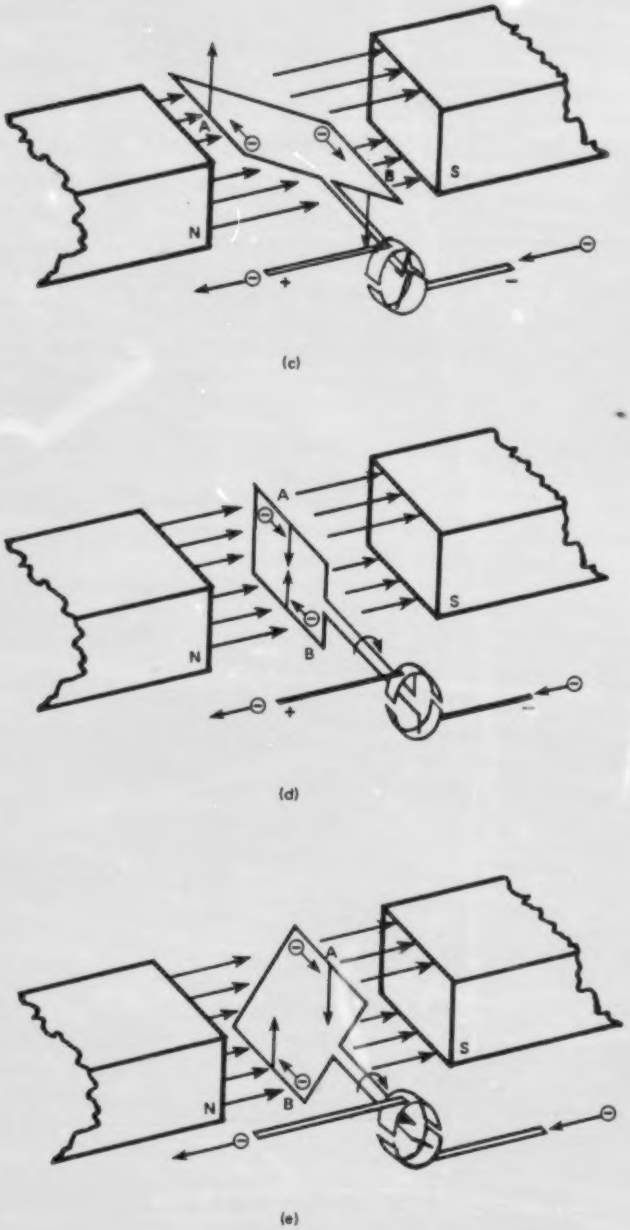
A short time later the armature has rotated to the position shown in Figure 3b. Since the forces acting on A and B depend only upon the current and upon the magnetic field strength, they have not changed. The torque acting on the armature has increased, however, because of the increase in the lever arm associated with each force. This configuration produces the maximum torque on the armature (Point "b" on Figure 4).

In Figure 3c, Segments A and B are still being pulled upward and downward, respectively. The torque is smaller because of a reduced lever arm (Point "c" on Figure 4), and the armature continues to rotate clockwise.

Just as the loop achieves the vertical orientation shown in Figure 3d, each brush slips to the opposite segment of the commutator ring, causing the direction of electron flow to be reversed. The forces on Segments A and B change direction, pulling A downward and pushing B upward. No torque acts on the loop, however, since the lever arm for each force is zero. (See Point "d" on Figure 4.)

Although no torque acts on the armature in Figure 3d, the momentum it already has built up allows it to coast into the position in Figure 3e, and the process repeats itself.

For ordinary magnetic fields and currents, the torque acting on a single loop of wire generally



(c)

(d)

(e)

FIG. 3 (CONTINUED FROM PREVIOUS PAGE)

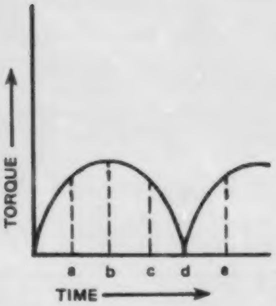


FIG. 4

is too small for practical application. Torque can be made larger by use of an armature that consists of many loops of wire, as shown in Figure 5. As current flows through the armature, a torque acts on each single loop. Each of these individual torques add to produce the total torque on the armature. Figure 6 shows the net torque acting on the motor of Figure 5 as a function of time. The maximum torque produced by the motor is determined by three factors:

- The number of loops in the armature.
- The strength of the magnetic field.
- The current flowing through the armature.

When any of these factors is increased, the torque increases.

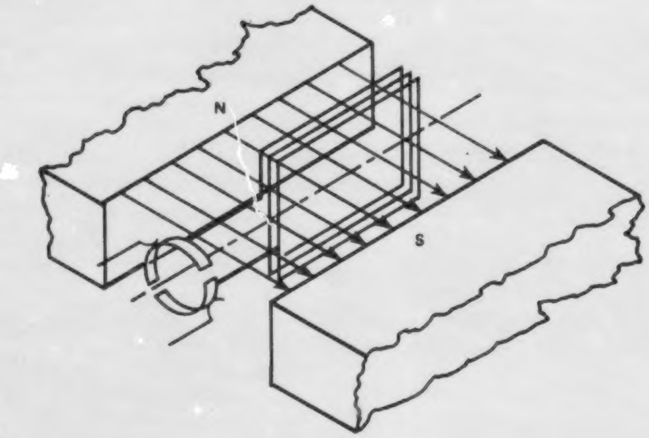


FIG. 5

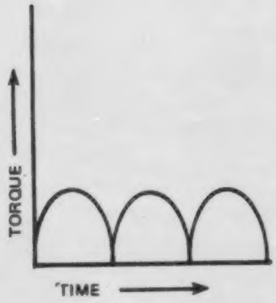


FIG. 6

Usually, permanent magnets can be used to provide the magnetic field in only very small motors. More powerful electric motors require a larger magnetic field than normally is available from permanent magnets. This problem is solved by the replacement of permanent magnets with electromagnets called "field coils."

Figure 7a shows a "series motor," so named because the field coils are connected in series with the armature. The same current flows through the field coils and the armature. Figure 7b shows a "shunt motor," in which the field coils are connected in parallel with the armature. Current through the field coils then is completely indepen-

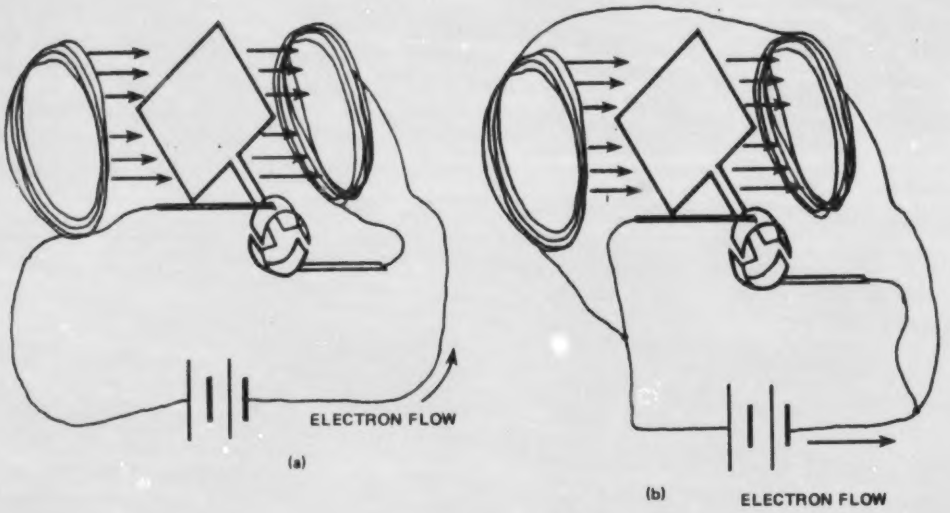


FIG. 7

dent of the armature current.

One problem with the motors discussed so far is that the torque acting on the armature varies greatly during a single revolution. This variation causes excessive wear on the bearings in the motor, and generally causes it to "run rough." Use of a modified armature, called a "multipole armature," causes the torque acting on it to be more nearly constant and to result in smoother operation.

A section view of the multipole armature is shown in Figure 8. This armature consists of many different coils of wire wrapped at equal intervals around the armature. The ends of each coil are connected to opposing segments of the commutator. The brushes are oriented in such a manner that they contact the commutator segments of each coil in the maximum torque orientation. The resulting torque is shown in Figure 9 as a function of time.

When an electric motor is rotating, the electrons in its armature windings actually undergo two distinctly different motions through the magnetic field. One motion, that of

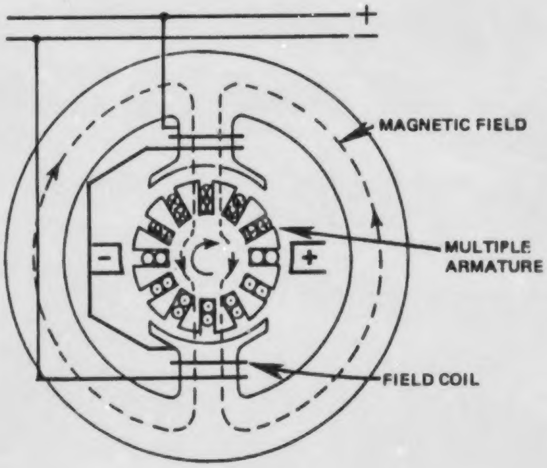


FIG. 8

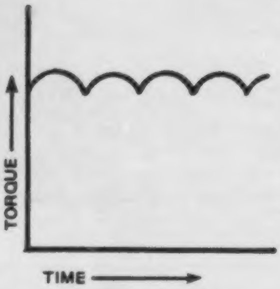


FIG. 9

the current flowing along the wire, already has been discussed in some detail in this module. This motion causes the magnetic forces that produce the torque on the motor; however, the electrons are carried around in a circular path by the spinning armature exactly as they would be in an electric generator. When the Left-hand Generator Rule is applied, an additional force is shown to be acting on the electrons. This force pushes against the flow of electrons in the armature, resulting in a voltage being produced across the armature. This voltage is opposite in polarity to the input voltage of the motor, and is called the "back emf." The back emf serves to decrease the armature current when the motor is spinning.

SUMMARY

The electric motor converts electrical energy into mechanical rotational energy. Its operation is made possible by the magnetic field inside the motor and by the forces exerted by the field upon conductors passing through it. The maximum torque developed by a particular motor depends upon (1) the number of loops in the armature, upon (2) the strength of the magnetic field, and upon (3) the current flowing through the armature.

LABORATORY

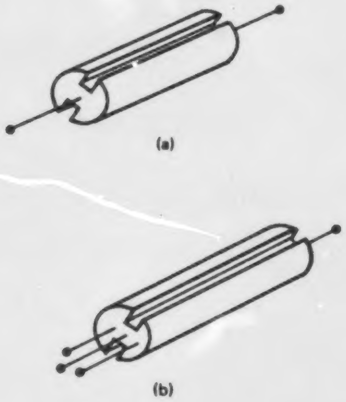
THE PERMANENT-MAGNET MOTOR

MATERIALS

Variable d.c. power supply, 0-6V, 5A, with meter to indicate output voltage.
Horseshoe magnet with north and south poles labeled.
Cork.
Razor blade.
Needle-nose pliers.
Eight straight pins.
Two thumb tacks.
Soft wooden base.
Twenty feet of #30 insulated copper wire.
Two strips of tin or copper foil 1/4" x 2".
Wire connectors.

PROCEDURE

1. Cut a groove in cork with razor blade, as shown in Figure 10a. Stick a pin in each end of cork to act as an axle.
2. Wrap wire in groove around cork to form armature of motor. Trim wire in order that coil starts and terminates at same end of cork. Leave about one inch of each end of the wire free. Strip ends by scraping with razor blade.
3. Make commutator by sticking two more pins in cork, as shown in Figure 10b. Wrap each end of armature coil around a different pin.


(a)

(b)

FIG. 10

4. Build mounting stand for motor by sticking remaining pins into wooden base, as in Figure 11a. Place armature in mounting stand; be sure that it is level and spins freely.
5. Construct brushes from strips of tin foil held in place by thumb tacks. Adjust brushes in such a manner that each makes simultaneous contact with commutator when armature is turned by hand. Contact between brushes and commutator should be light, in order that armature rotation will not be

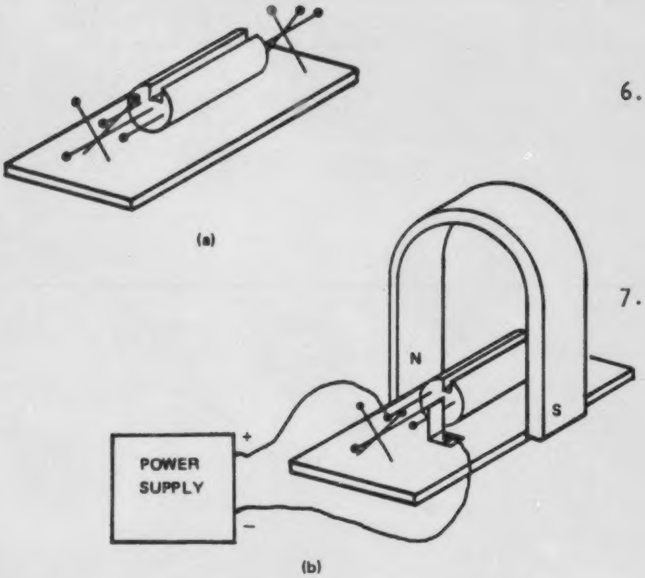


FIG. 11

impeded.

6. Run a separate wire connector from each pole of power supply to one of the brushes, as shown in Figure 11b. Place horseshoe magnet over armature, as shown. The motor is now complete.

7. Label Figure 11b with the following information:

   - Direction of magnetic field.
   - Direction of electron flow through the armature.
   - Forces acting on the horizontal segments of wire in the armature.
   - Direction of armature rotation.

8. Turn on power supply, and set it to about 1.5 V. Spin motor gently to get it started. Record direction of rotation in Data Table 1.

9. Reverse poles of magnet so that direction of magnetic field is opposite its original direction. (Turn off power supply before making any changes.) Record direction in which the armature spins. Make a sketch similar to Figure 11b to justify this observation.

10. Return the magnet to its original orientation, and reverse leads on power supply. Record direction in which motor now spins, and justify with a sketch.

11. Record and justify with a sketch the effect of reversing both the magnet and power supply leads from their original positions.

THE SERIES MOTOR

MATERIALS

Pin and cork motor built in previous laboratory exercise.
Variable d.c. power supply, 0-6V, 5A, with meter to indicate output voltage.
Two nails 4" in length.
Two blocks of wood.

Seven feet of #20 insulated wire.

PROCEDURE

1. Drive one nail approximately one inch into each block of wood. When blocks are laid on their sides, each nail should be parallel to table top, at approximately equal heights above it. Heads of nails should be at about the same height as axle of motor.

2. Strip about 1/2 inch of insulation from each end of #20 wire. Refer to Figure 12. Starting about 3 inches from one end, wrap about three feet of wire around one nail. Starting about three inches from opposite end, neatly wrap three feet of wire around other nail. Be sure that wire is wound in same direction around each nail. When you are finished, wrapped nails should look like those in Figure 2a. These are the field coils.

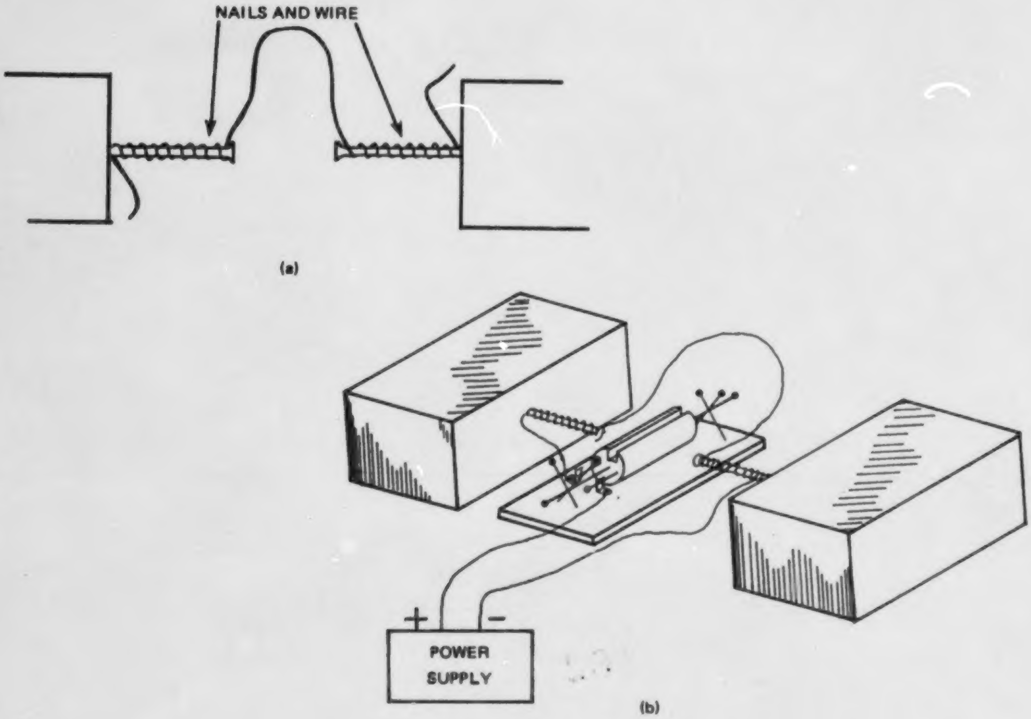


FIG. 12

3. Connect field coils, motor, and power supply, as shown in Figure 12b. Wires may be attached to each other by twisting them together. Place field coils opposite each other, with armature directly between them. Heads of nails should be as close to one another as possible without interfering with spinning armature.

4. Turn on power supply, and adjust it to approximately 3V. Start motor spinning, and record direction of spin in Data Table 2.

5. Determine and record effect of the following changes on direction of rotation of armature:

    - Changing each field coil to the opposite side of armature.
    - Reversing power supply leads.
    - Reversing both field coil positions and power supply leads.

6. What are the basic difference between the permanent-magnet, series, and shunt motors?

DATA TABLE 1

Original spin direction: ____________________

Spin direction with only magnet reversed: ____________________

Spin direction with only power supply leads reversed: ____________________

Spin direction with both magnet and leads reversed: ____________________

DATA TABLE 2

Original spin direction: ____________________

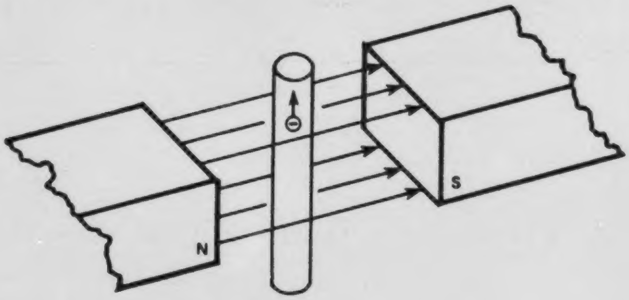
Spin direction with field coils reversed: ____________________

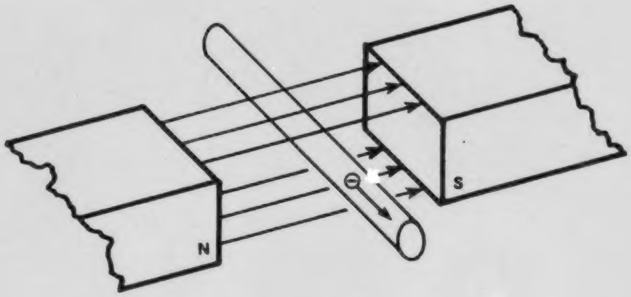
Spin direction with power supply leads reversed: ____________________

Spin direction with field coils and power supply leads reversed: ____________________

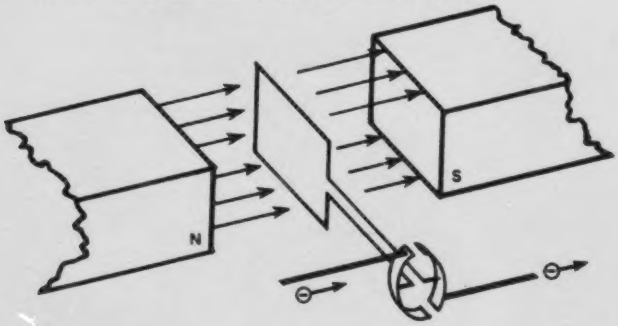
SELF CHECK

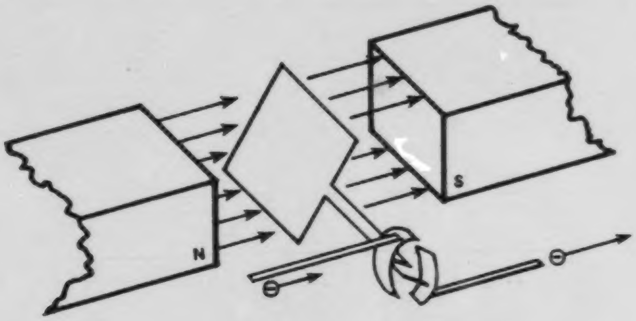
1. Determine the direction of the force on the current-carrying wires in the following diagrams. The direction of electron flow is indicated in each case.

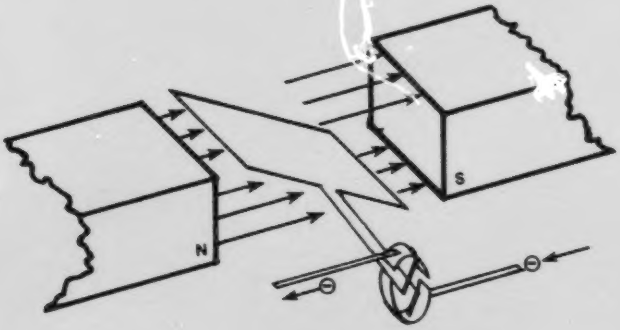




2. Indicate the direction of rotation of the current-carrying loops in the following diagrams. The direction of electron flow is indicated with an arrow in each case.







3. Sketch a permanent-magnet electric motor. Label the magnets, armature, commutator, and brushes; and explain the function of each.

4. List the three factors that determine the maximum torque produced by a motor.

5. State the difference between the permanent-magnet, series, and shunt motors.

6. According to Figure 6, if the torque acting on a motor reaches its first maximum at $t = 3$ seconds, when will the torque reach its second maximum?

7. Certain functions can be most easily expressed by two different functions, each of which is valid over a different subset of the domain (abscissa). Assuming the curve drawn in Figure 6 is in general shape a sine curve, write two expressions for torque as a function of time which describe the curve. Define the time from $t = 0$ until the time that the torque first reached zero again to be one-half of a period ($\tau/2$) and write one expression which is good for times

$$0 \leq t \leq \frac{\tau}{2}, \quad \tau \leq t \leq \frac{3\tau}{2}, \quad 2\tau \leq t \leq \frac{5\tau}{2}, \text{ etc.}$$

and a second expression good for times

$$\frac{\tau}{2} \leq t \leq \tau, \quad \frac{3\tau}{2} \leq t \leq 2\tau, \quad \frac{5\tau}{2} \leq t \leq 3\tau, \text{ etc.}$$

8. One end of a metal rod experiences a torque of 3 N·m at a distance of 1 m from the point at which the rod is pivoted. If the force on the rod is perpendicular to the length of the rod, find the magnitude of the force on the end of the rod.

9E4

SOLENOIDS AND RELAYS

Electromechanical devices are energy converters that use both electrical and mechanical energy. In a solenoid, an electrical current flows through a coil of wire to produce a magnetic field. This magnetic field causes an iron plunger to be magnetized and drawn into the center of the coil, converting electrical energy into mechanical energy. The resulting mechanical motion can be used to operate switches, valves or other mechanical devices. When a solenoid is used to operate a switch, the entire device is called a relay. This module discusses the operation and applications of the solenoid and the relay. In the laboratory, the student will identify relay terminals and will operate starter and control circuits.

OBJECTIVES

Upon completion of this module, the student should be able to:

- List and explain the operation of the three components of a solenoid.
- Describe how the plunger of a solenoid is drawn into the coil form by a magnetic field.
- State the purpose of a relay.
- Draw and label a diagram showing a single-pole, single-throw relay in a control circuit.
- Draw and label diagrams showing the contacts of the following relays:

    Single-pole single-throw.
    Single-pole double-throw.

Double-pole single-throw.
Double-pole double-throw.

- List three electrical devices that employ relays.
- Given the appropriate equipment, identify relay terminals and operate an automobile starter circuit and an electric motor control circuit.

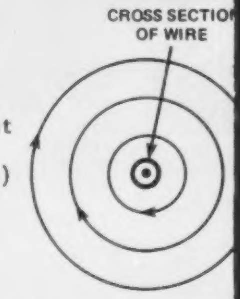
SOLENOIDS

The Danish scientist Hans Christian Oersted discovered in 1820 that an electric current flowing in a wire causes a magnetic field around the wire. The production and control of a magnetic field by an electric current is the basis of the solenoid.

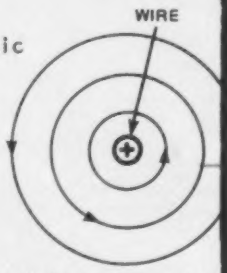
Figure 1 shows the magnetic field produced by a long, straight wire carrying a current. In Figure 1a, electron current flows toward the reader out of the plane of the page (indicated by a dot) and sets up the resulting circular magnetic field, pointing in a clockwise direction. In Figure 1b, the current direction is reversed to flow into the page (indicated by a cross), and the direction of the magnetic field also is reversed.

In Figure 2, two current-carrying conductors are placed close together. If both currents flow in the same direction (into the page in this case), the magnetic fields combine, as shown, to produce a stronger field. Additional conductors with the same current direction would further increase the strength of the magnetic field.

If the wire is coiled as shown in Figure 3, a stronger magnetic field within the coil results. The arrows in the figure show the direction of the magnetic field around the wire at all points. The fields in Figure 3 always point downward inside the coil and upward outside. These small fields produce the total magnetic field of the coil shown by dotted field lines. This magnetic field produces mechanical motion in a solenoid.



(a) CURRENT OUT OF PAG



(b) CURRENT INTO PAG

FIG. 1

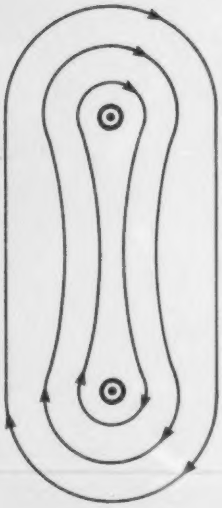


FIG. 2

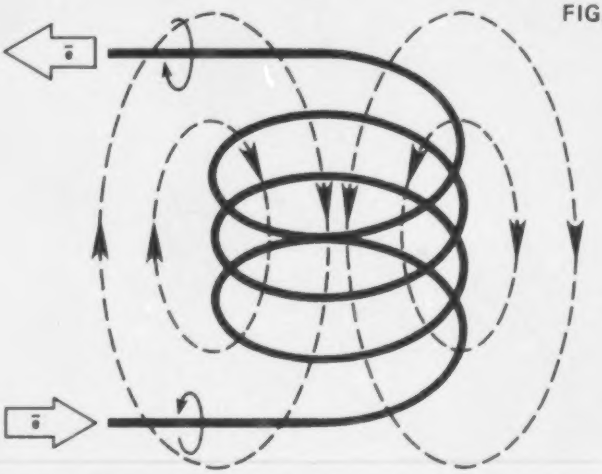


FIG. 3

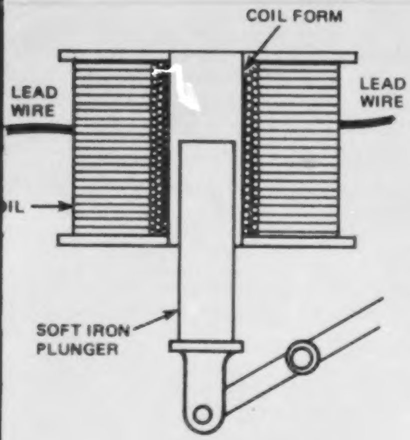


FIG. 4

Figure 4 shows the three components of a solenoid. A coil of wire is wrapped around a hollow coil form made of a non-magnetic material. The plunger is a cylinder of ferro-magnetic material (usually iron) that slides easily inside the coil form. The plunger is situated partially inside the coil form when the coil is not energized.

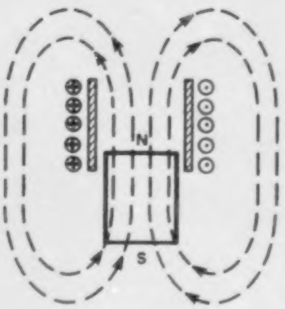


FIG. 5

When a current begins to flow in the coil (Figure 5), the magnetic field around the coil is established. The coil behaves as a magnet with N pole at the top and S pole at the bottom. The field is distorted by the presence of the iron plunger. It passes through the iron and causes it to be magnetized as shown. The north pole of the plunger is pulled upward by the magnetic field of the solenoid, and the south pole is pushed downward. Because the north pole is in a stronger part of the field than the south pole, the upward force is greater than the downward force; and the plunger moves upward until it is centered in the coil form (Figure 6). Upward motion of the plunger can be used to operate valves, switches or other mechanical devices. When the current is interrupted, the magnetic field disappears and the plunger is returned to its original position by gravity or by a spring. Solenoids are used in a variety of applications including liquid and gas flow valves, door bells and electric locks for automobile doors.

RELAYS

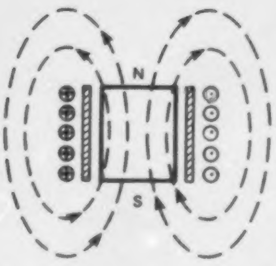


FIG. 6

The most common application of the solenoid is in a device called a relay (Figure 7). A solenoid is used in a relay to operate a switch, allowing a small voltage and current to be used to control larger voltages or currents. In Figure 7 the control switch is open and no current flows in the coil of the relay. The plunger is held down by a spring, and the relay switch is open, which prevents current flow in the controlled circuit. When the control switch is closed (Figure 8), a current flows in the relay coil, pulling the plunger into the coil form. This stretches the spring and closes the relay switch, completing the controlled circuit. When the control switch is opened again, the magnetic field disappears; and the spring pulls the plunger downward to reopen the relay switch.

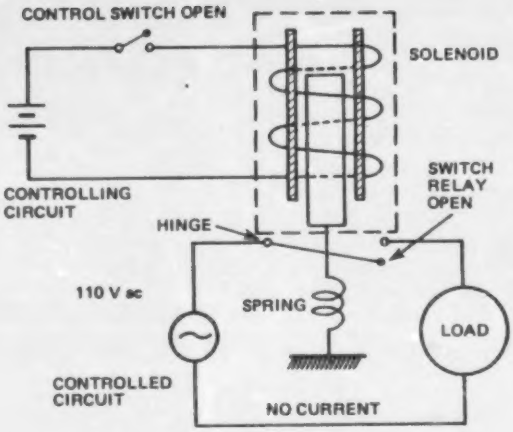

CONTROL SWITCH OPEN
SOLENOID
SWITCH RELAY OPEN
CONTROLLING CIRCUIT
HINGE
110 V ac
SPRING
LOAD
CONTROLLED CIRCUIT
NO CURRENT

FIG. 7

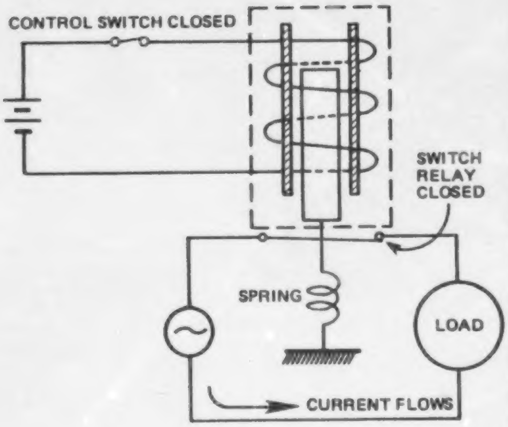
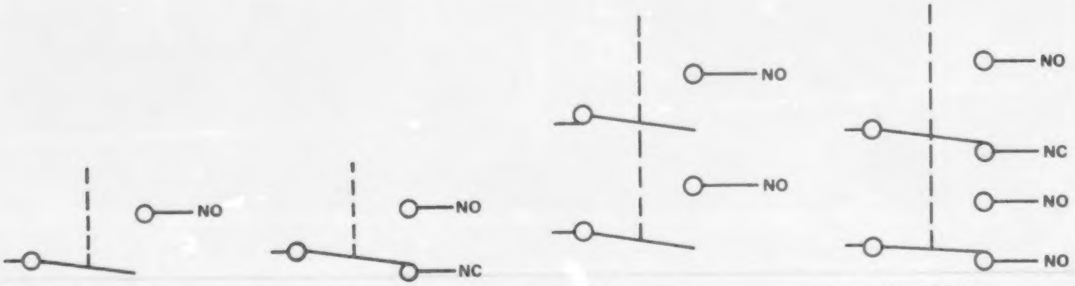

CONTROL SWITCH CLOSED
SWITCH RELAY CLOSED
SPRING
LOAD
CURRENT FLOWS

FIG. 8

The relay shown in Figure 7 is called a single-pole, single-throw (SPST) relay, which means that it contains a single switch that completes one circuit when the control switch is closed (thrown). In many cases, control of more than one circuit with one relay is desirable. This control is accomplished by adding more switches or by allowing a relay to break one circuit when it completes another. Switch contacts that are closed when the relay coil is not energized are called normally-closed (NC) contacts. Those that are open when the coil is not energized are called normally-open (NO) contacts. Figure 9 shows the switch contact arrangements for four common types of relays.


NO
NO
NC
NO
NO
NO
NC
NO
NO

a. SINGLE POLE SINGLE THROW–SPST

b. SINGLE POLE DOUBLE THROW–SPDT

c. DOUBLE POLE SINGLE THROW–DPST

d. DOUBLE POLE DOUBLE THROW–DPDT

FIG. 9

Relays are available in a wide variety of sizes and designs. Some contain many more switches than shown by the examples in Figure 9 for controlling more circuits. Some are designed for d.c. coil current and others for a.c. coil current. The following specifications often are given on a relay, and are important in selecting the proper relay for a specific job:

- Coil voltage.
- Coil current.
- Number and type of switches.
- Switch current.
- Switch voltage.

Relays are found in most electrical systems in which large currents must be controlled. The following are three common examples:

- In an automobile, the starter switch operates a relay that controls the starter motor.
- In an electric hot water heater, a temperature sensing device controls a relay that controls the current to the heating elements.
- In an air conditioner, the thermostat operates a relay that controls the current to the compressor and fan motors.

SUMMARY

The solenoid coverts electrical energy into mechanical energy. It consists of a cylinder-shaped coil of wire through which current is passed and of an iron core (plunger) which is moved by the electromagnetic field produced by the coil. The relay contains a solenoid attached to a switch. Solenoids and realys are used widely in fluid flow control and in a variety of switching operations.

IDENTIFYING RELAY TERMINALS

MATERIALS

12 V d.c. relay SPDT.
VOM.

PROCEDURE

1. Examine relay, and determine which terminals are connected to movable part of relay switch, which are connected to fixed switch contacts and which are connected to relay coil.

2. Set VOM to ohms x 1 scale, and measure resistance of relay coil. Coil resistance varies with relay design but should be between 15 Ω and 200 Ω.

3. Attach one VOM lead wire to terminal connected to movable part of relay switch. Connect other VOM lead to one of terminals connected to a fixed switch contact, and note resistance measured.

4. Repeat with other terminal.

5. From the data determined, identify terminals connected to fixed switch contacts as normally open (infinite resistance reading) or normally closed (zero resistance reading).

AUTOMOBILE STARTER CIRCUIT

MATERIALS

12 V d.c. relay SPDT.
12 V d.c. power supply or 12 V battery.
12 V lamp (to simulate starter motor).
Normally-open push-button switch.
Connecting wires.

PROCEDURE

1. Connect circuit as shown in Figure 10.

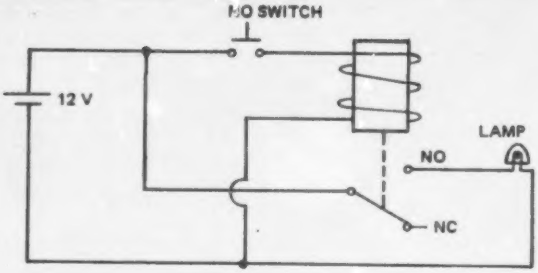


FIG. 10

2. Alternately press and release push button switch while observing relay and lamp operation.

3. Describe circuit operation, and explain why the relay closes.

ELECTRIC MOTOR CONTROL CIRCUIT

MATERIALS

12 V d.c. relay SPDT.
12 V d.c. power supply or 12 V battery.
Two 12 V lamps.
Normally-open push-button switch.
Normally-closed push-button switch.
Connecting wire.
SPST toggle switch.

PROCEDURE

1. Connect circuit as shown in Figure 11.

2. Turn on power toggle switch.

3. Press and release "on" push-button switch.

4. Press and release "off" push-button switch.

5. Observe circuit operation.

6. If this circuit were attached to a motor as shown in Figure 11, would the motor operate when step 2 was completed, or when step 3 or 4 was completed? Explain.

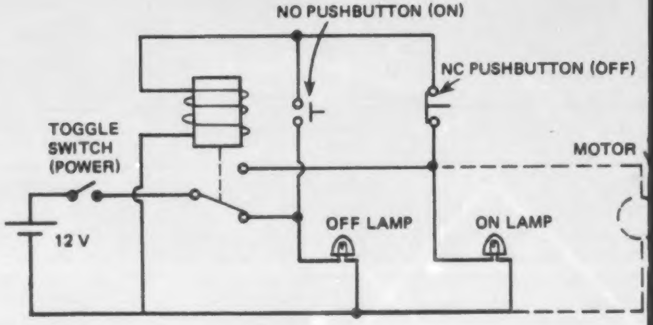
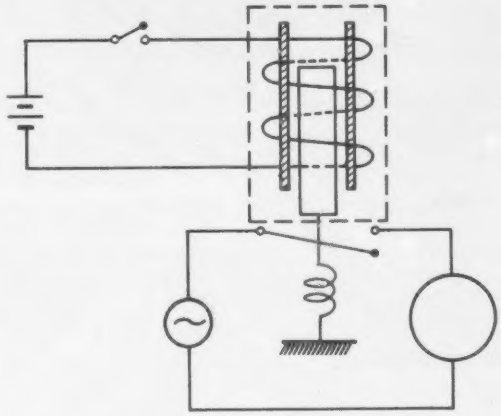
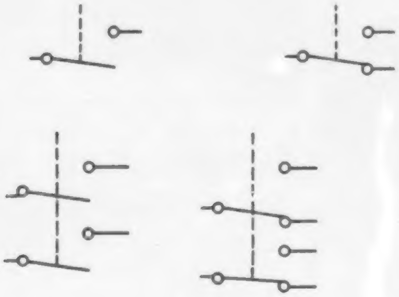

NO PUSHBUTTON (ON)
NC PUSHBUTTON (OFF)
TOGGLE SWITCH (POWER)
MOTOR
12 V
OFF LAMP
ON LAMP

FIG. 11

SELF CHECK

1. List the three components of a solenoid.
2. Explain how the plunger of a solenoid is drawn into the coil form by the magnetic field.
3. What is the purpose of a relay in an electrical circuit?
4. Label the diagram at the right.
5. Identify each of the relay contact configurations shown below.



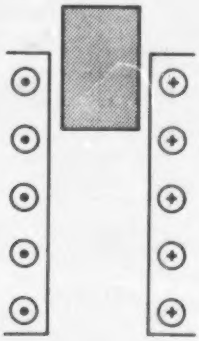


6. Name three devices using relays.

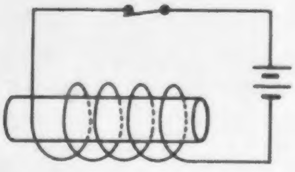
7. Indicate the direction of a magnetic field that would be produced by an electron current flowing as shown below.

DIRECTION OF CURRENT FLOW

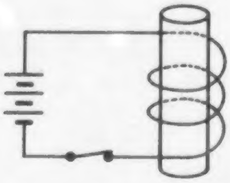
8. Draw the magnetic field lines of the following and indicate the direction of movement of the iron plunger in the coil shown below.



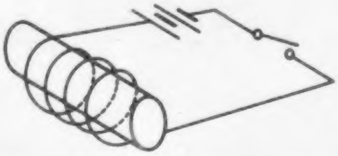
9. What direction do the magnetic field lines go inside the coil form for each of the following coils according to the flow of current.



a.



b.



c.

9E5

PHOTOVOLTAIC MATERIALS

Photovoltaic materials are semiconductors; specifically, they are junctions between a P-type semiconductor and an N-type semiconductor. When light — for example, sunlight — strikes such materials, electrical current is generated. Thus, photovoltaic materials can be used as detectors of light and also as energy converters that convert light directly into electricity.

This module examines the operation and construction of photovoltaic materials. In the laboratory, the student will use a photovoltaic device to measure efficiency of conversion of light energy into electrical energy.

OBJECTIVES

Upon completion of this module, the student should be able to:

- Describe the operation of photovoltaic devices, including energy bands in semiconductors, P-N junctions, generation of free electrical carriers by absorption of light energy near the junction leading to electrical current flow; name at least one useful photovoltaic semiconductor material.
- Describe at least two applications of photovoltaic materials.
- Given the appropriate equipment, use a photovoltaic device to measure the efficiency of conversion of light energy into electrical energy.

DISCUSSION

An excellent example of an energy converter is the photovoltaic transducer. When energy in the form of light is incident on a photovoltaic material, it is converted directly into electrical energy. Because of this, the photovoltaic transducer has great utility for quantitative measurement of light energy and also for direct generation of electrical energy from light energy such as sunlight.

To understand operation of photovoltaic transducer, we must refer to some of the concepts related to semiconductors, as described in Module 6E5. Most materials used in photovoltaic transducers are semiconductors; therefore, some of the concepts related to semiconductors must be reviewed.

HOW SEMICONDUCTORS WORK

N-TYPE JUNCTIONS

As a brief review, remember that a semiconductor is a material with values of electrical conductivity intermediate between those of good metallic conductors, such as copper, and electrical insulators, such as rubber. Examples of semiconductors include materials such as silicon and germanium. The "energy band structure" of semiconductors is important. Current can be carried through a semiconductor by the motion of electrons; however, free electrons in a semiconductor are few in number compared to those in a metal, so that electrical conductivity is much lower. In addition, in a metal the free electrons have a continuous range of energy levels available to them. In a semiconductor, on the other hand, there are bands of allowed energy, separated by a band gap, a range of energies not available to the electrons (Figure 1). Free electrons in the conduction band are free to move under action of an electric field and contribute to electrical conductivity. If the dominant contribution to electrical conduction is free electrons, the material is termed an N-type semiconductor.

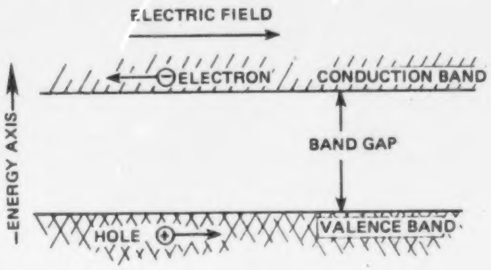


FIG. 1

P-TYPE JUNCTIONS

Another contribution to electrical conduction can be through missing electrons in the valence band. As electrons from neighboring sites move to fill the vacancy, the position of the missing electron may move. The movable vacancy appears similar to motion of positive charge through the material. The movable vacancy is called a "hole," and it acts like a positive charge carrier. Materials in which the dominant contribution to electrical conduction is the motion of the holes are termed "P-type" semiconductors. Semiconductors can be made P-type or N-type when doped in manufacture with small amounts of appropriate impurity elements.

P-N TYPE JUNCTIONS

A P-N junction is a region in which there is an interface between P-type and N-type material, of very small dimensions, typically of the order of a few micrometers. Such a junction is called a "P-N junction," which is the important junction for photovoltaic effects.

Most modern photovoltaic devices employ a P-N junction in a semiconductor. The photovoltaic effect also appears at a function between a metal and a semiconductor; in fact, many early photovoltaic devices employed such a metal-semiconductor junction. A common type is the junction between iron and the semiconductor selenium. Such devices found great utility as photovoltaic cells and were used to turn on street lights automatically when darkness fell. Since the applications of optical energy measurement and energy conversion usually employ semiconductor photovoltaic materials, the emphasis in this module will be upon semiconductors.

The photovoltaic effect depends upon distortion of the energy band structure through a P-N junction (Figure 2). Height of the conduction band and valence band change through the junction region from N-type material to P-type material. Thus, an electron in the P-type material will move along the bottom of the conduction band, as shown. Similarly, a hole near the top of the valence band in the N-type materials will also move as shown. Thus, the junction acts very similarly to a built-in electrical field, causing the motion of the electrical charge.

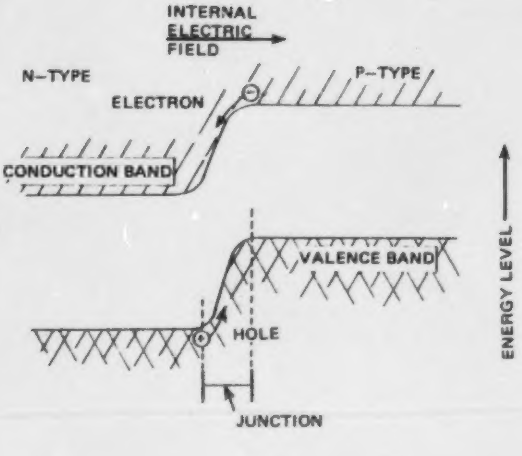


FIG. 2

Thus, when P-type and N-type semiconductors form a junction, a voltage is produced across the junction which will separate holes and electrons. However, there usually are no free holes and electrons in in the immediate vicinity of the junction to allow electrical current to flow. If, however, light is incident upon the junction and is absorbed, the light energy will produce free holes and electrons, which then are capable of moving under the action of the electric field (Figure 3).

Absorption of light energy in the vicinity of the junction thus creates current carriers, which are separated by the built-in electric field. If electrical contacts are attached at the surface on the two sides and connected through a load resistor as shown, there will be a net flow of electrical current through the load resistor and, therefore, generation of electrical power.

This generation is the photovoltaic effect — direct generation of electrical energy by absorption of light energy in the vicinity of a P-N junction in a semiconductor. No external voltage must be applied to produce the photovoltaic effect.

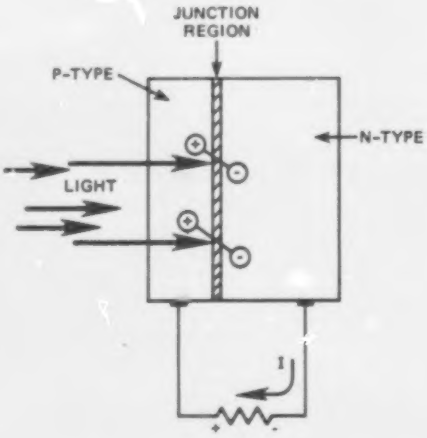


FIG. 3

OTHER PHOTOSENSITIVE DEVICES

There are several other optical-electrical effects that rely upon absorption of optical energy and through which current may be generated in an external load resistor. Some examples are the "photoemissive effect" and the "photoconductive effect." Energy converters and optical sensors based upon these effects also have been developed. Both of these effects require an external voltage applied to the materials.

OPERATIONAL CHARACTERISTICS OF SILICON PHOTOVOLTAIC DEVICES

CONSTRUCTION AND ELECTRICAL CIRCUITS

The circuit employed for photovoltaic devices is illustrated in Figure 4. The light (arrows) is incident on the photodetector, shown schematically as a diode. Current I flows as shown, and the voltage is generated across the load resistor R_L. Voltage is given by Ohm's law (Equation 1).

$$V = IR_L \qquad (1)$$

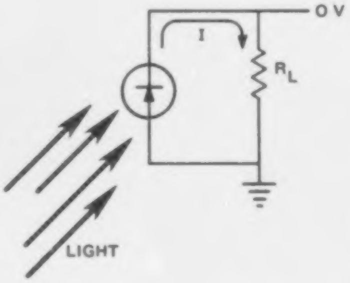


FIG. 4

where
V = voltage.
I = current.
R_L = load resistor.

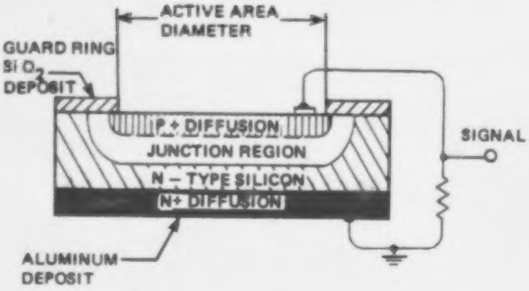


FIG. 5

Structure of a typical silicon photovoltaic device is displayed in Figure 5. A piece of N-type silicon is used as the starting material. Impurities are diffused into one side to form a P-type region and a junction. Other materials can be diffused into the opposite side to produce a heavily-doped N-type region for ease in making an electrical contact. A guarding structure of SiO_2 is deposited on the surface to reduce stray current leakage. Electrical contacts (aluminum) are deposited on the surface. Fabrication of such devices has become a well-developed technology.

RESPONSIVITY

The response of the P-N junction to light is represented by the response curve in Figure 6. For relatively low values of the load resistor, response — that is, externally generated current — is approximately a linear function of the incident light energy. For large values of the load resistor, the response will saturate, as indicated. Response will saturate when the voltage drop across the load resistor becomes equal to the value of built-in voltage at the junction. Thus, the voltage drop has a maximum value, equal to the product of the current and the resistance. Therefore, for larger values of resistance, there is a maximum limiting current that can be generated.

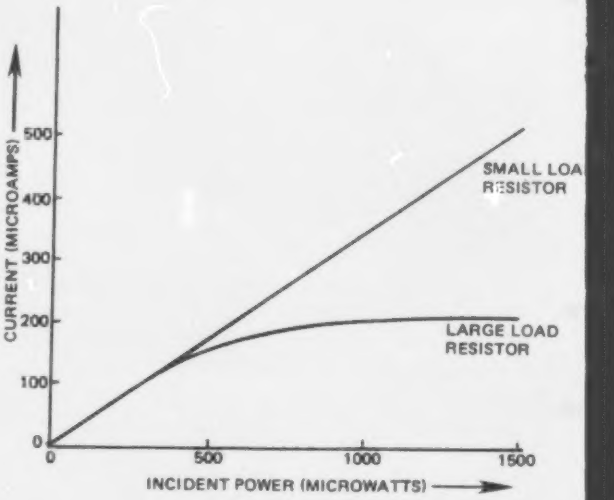


FIG. 6

The photovoltaic effect also can be used in the form of a sensor that generates a measured voltage across an open circuit. The sensor is less important as an energy converter than some other applications because it requires the development of electrical power and, therefore, the flow of electrical current, in the external circuit.

Figure 6, defines an important parameter of the photovoltaic detector — namely, "responsivity," which is the amount of current that flows in the external circuit divided by the optical input power. It has units of current per unit power and is given by the quotient of the current divided by the power. For the example shown in Figure 6, responsivity is equal to 210 microamps divided by 600 microwatts, or 0.35 microamps per microwatt. This is a typical value characteristic of modern photovoltaic cells based upon silicon P-N junctions.

EFFICIENCY

Current generated by absorption of light delivering power absorbed by the junction is given by Equation 2.

$$I = RP \tag{2}$$

where

R = responsivity.
P = power in watts.

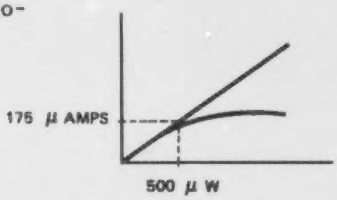
Since electrical power generated will be IV, with V, voltage, given by IR_L, the efficiency (ratio of electrical power to light power) can be expressed as in Equation 3.

$$\text{Efficiency} = \frac{IV}{P} = \frac{I^2R_L}{P} = \frac{R^2P^2R_L}{P} = R^2PR_L \tag{3}$$

Notice that efficiency cannot be increased very much by an increase of R_L, because of saturation effects shown in Figure 6 when a large resistor is used. A calculation using Figure 6 is given in Example A.

EXAMPLE A: CALCULATION OF CURRENT AND VOLTAGE IN A PHOTOVOLTAIC CELL

Given: Light is incident upon a photovoltaic cell that has characteristics shown in Figure 6. The load resistor, as in Figure 5, is 100 Ω. Optical power is 500 microwatts.



Find: (a) Current I and (b) voltage V.

Solution: (a) The resistor is a relatively small one; therefore, from the upper curve in Figure 6, current I is read as 175 microamps at 500 microwatts optical power.

(b) Use Ohm's law (Equation 1):

$$V = IR_L$$
$$= (175 \times 10^{-6})(100)$$
$$V = 1.75 \times 10^{-2} \text{ volts.}$$

SPECTRAL RESPONSE

Spectral response of silicon is plotted in Figure 7. This graph shows response of silicon photovoltaic devices as a function of the wavelength of the incident light. (A full discussion of the concept of wavelength of light will be developed in Concept Module 13-0.) The reason for the shape of this curve is beyond the scope of this module. We simply point out that silicon detectors are most useful in the visible portion of the spectrum (0.4 to 0.7 micrometers), and in the near-infrared portion of the spectrum (0.7 to 0.9 micrometers). Other materials exhibit a strong response farther into the infrared and, therefore, have applicability as detectors for longer-wavelength infrared radiation, to which silicon detectors would not respond.

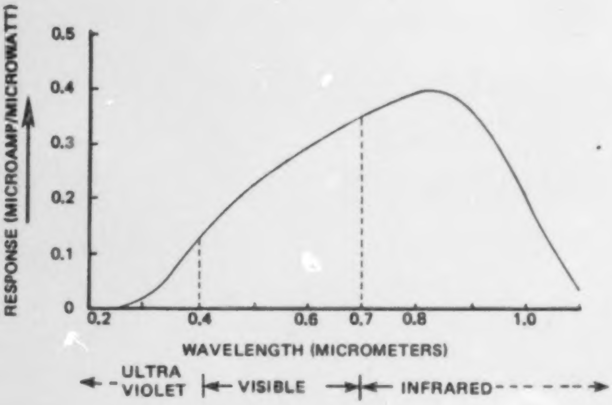


FIG. 7

OTHER DEVICES USING DIFFERENT SEMICONDUCTOR MATERIALS

A number of different types of semiconductors have been employed, including silicon, gallium arsenide, indium phosphide, cadmium telluride, gallium phosphide, and cadmium sulfide. Some of the other materials have a higher theoretical efficiency than silicon, which has a maximum theoretical efficiency of approximately 20%. However, for large-scale applications, where low cost is a factor, silicon now is the most practical candidate. Some of the other materials can be used as detectors of infrared radiation.

APPLICATIONS

SENSORS

As previously mentioned, photovoltaic energy converters are used to turn street lights on and off automatically. They also are commonly employed as quantitative detectors of the amount of optical radiation. Thus, they have application in camera light meters, photometers (used to measure illumination levels), colorimeters (used to measure color values of paints, fabrics, etc.), and infrared detectors (used in night-vision devices and in medical sensors). They can be used as electric eyes (which could be employed to open doors automatically, or as burglar alarms), as flame sensors (fire alarms and furnace controls) and as liquid-level sensors.

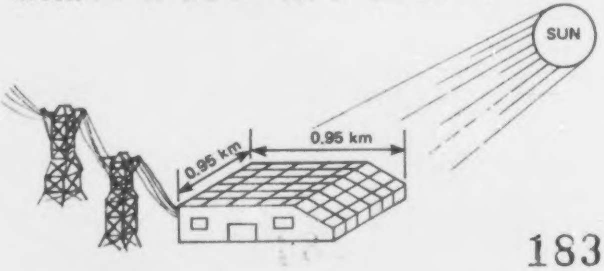
ELECTRICAL POWER GENERATION

As energy converters, used for direct generation of electrical energy from sunlight, photovoltaic converters have been employed for a number of years on satellite systems to provide electrical energy for satellite operation. These are the solar cells that have been employed on many satellites and space probes. Such cells have not found much application in direct generation of electricity from sunlight on the earth's surface, because of the relatively high cost of the material. High cost has not been a great drawback for space applications because total areas involved are small and because the very specialized application will tolerate a higher cost per unit of electrical power. For generation of larger amounts of electrical power on the earth's surface, however, the cost will have to be reduced considerably. If the cost of silicon photovoltaic cells could be reduced to the approximate region of $0.50 per watt of electrical generating capacity, they could be competitive on the earth's surface for direct generation of electricity from sunlight. By the mid-1980's, the refinement of silicon technology probably will allow this cost reduction; therefore, within the 1980's the direct use of photovoltaic energy converters to produce electricity on the earth's surface may be a reality.

Example B shows the calculation for the size of a photovoltaic cell unit used in a generating station.

EXAMPLE B: CALCULATION OF THE SIZE OF A SOLAR CELL UNIT FOR A GENERATING STATION

Given: A solar electrical generating station based upon photovoltaic energy converters uses silicon solar cells having an efficiency of 10%. Sunlight delivers 0.11 watts/cm^2 to the surface of the earth.



Find: Size of the area that must be covered by the solar cells for a 100-megawatt generating station.

Solution: The solar power to be collected is 1000×10^6 watts (100 megawatts divided by 10% efficiency). This much power will require an area of $1000 \times 10^6 \div 0.11 \simeq 9.1 \times 10^9$ cm^2, which would require a square area about 9.5×10^4 cm (0.95 km) on a side to be covered by the silicon solar cells. Notice that the station will be operative only when the sun is shining.

SUMMARY

Photovoltaic materials are semiconductor devices (P or N) that convert light energy into electrical energy. Some combinations of semiconductor materials deposited upon metal plates can generate voltage at the junctions when exposed to light radiation. Voltage is generated at junctions (barrier layers) between the two types of materials when radiant energy is absorbed. This process is termed the "photovoltaic effect."

LABORATORY

EFFICIENCY OF A PHOTOVOLTAIC DEVICE

MATERIALS

Photovoltaic detector (e.g., United Detector Technology Model PIN-10DP).
Helium-neon laser (about 1-2 mW, e.g., Spectra-Physics Model 132) or other low-power light source.
Resistors - 100 Ω and 10,000 Ω.
Milliammeters.
Wire for electrical connections.
Laser power meter (e.g., Spectra-Physics Model 401C).
4 — Neutral density filters (e.g., Kodak Wratten filters having optical density of approximately 0.3).
Graph paper.

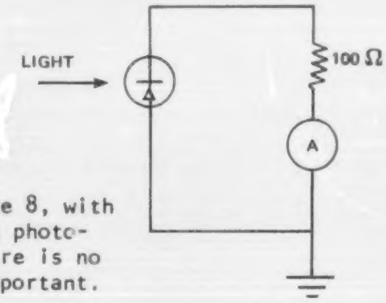


FIG. 8

PROCEDURE

1. Connect photovoltaic device as illustrated in Figure 8, with the 100-Ω load resistor and ammeter in series. The photovoltaic device is shown as a diode; but because there is no externally-applied voltage, polarity will not be important.

2. CAUTION: Do not look directly into laser beam. Do not allow beam to reflect from a shiny object into eyes. Turn on the HeNe laser. Direct beam into laser power meter, and measure power. Insert neutral-density filters between laser and power meter. Measure power transmitted through different numbers of neutral-density filters. Enter values in Data Table.

3. Remove filters and power meter. Direct beam to strike photovoltaic detector. Make sure that beam falls entirely on active area of detector. Measure current on milliammeter.

4. Add neutral-density filters one by one, repeating Step 3, and measure current in each case. Enter values in Data Table.

5. Replace 100-Ω resistor with 10,000-Ω resistor. Repeat procedure in Steps 3 and 4.

6. Plot results obtained in Steps 3, 4, and 5 on graph paper. Plot measured current versus incident light power for the

two values of load resistance.

7. From the 100-Ω resistor data, calculate responsivity (microamps/microwatt).

8. From value in Data Table for the 100-Ω resistor and no neutral-density filters, calculate efficiency of conversion of optical energy into electrical energy for this experiment. Remember, electrical power (in watts) is equal to square of current (in amperes) times load resistance (in ohms).

9. Why is the shape of the curve different for the two different resistors?

DATA TABLE

No. of Neutral Density Filters	Laser Power Transmitted	Current (100 Ω Resistor)	Current (10,000 Ω Resistor)
0			
1			
2			
3			
4			

SELF CHECK

1. Describe operation of photovoltaic devices.

2. Describe at least two applications of photovoltaic materials.

3. Suppose a photovoltaic detector has a responsivity of 0.5 microamps per microwatt and that 10 milliwatts of light energy are absorbed by the device. What is the resulting electrical current if the device is connected across a load resistor of small resistance?

4. Given results of problem 3, with a load resistance of 120 Ω, what is the energy efficiency of conversion of light energy into electrical energy?

5. What is the responsivity of a junction which generates 40 microamps of current when absorbing 100 microwatts of power?

6. How large should a load resistor be if a current of 0.1 mA is to produce a voltage of 1 V?

7. How much power does sunlight deliver, per square centimenter, to a photovoltaic energy converter that is 6% efficient and produces a 110 mA current across a 1000-Ω resistor if the sunlight is collected on a 800 cm^2 area?

9T2

THERMOELECTRIC GENERATORS

THERMOELECTRIC GENERATORS

This module explores thermoelectricity, the direct conversion of heat energy into electrical energy (or the reverse) by means of three related physical phenomena — the Seebeck effect, the Peltier effect, and the Thomson effect. Thermoelectric effects are exhibited by both metallic conductors and by semiconductors. The magnitude of the effects is larger in semiconductors; therefore, semiconductors generally are used as practical thermoelectric generators, which, in turn, are used to produce electricity directly from heat. In the laboratory, the student will set up a simple thermoelectric apparatus and will measure the thermoelectric voltage generated for both metallic conductors and for semiconductors.

OBJECTIVES

- Upon completion of this module, the student should be able to:
- Describe the Seebeck effect, the Peltier effect, and the Thomson effect.
- Name two general types of thermoelectric materials, and give at least one specific example of each.
- Name at least four applications of thermoelectricity.
- Given the necessary information about materials, calculate the thermal EMF generated in a specific arrangement.
- Given the necessary information, calculate the efficiency of a thermoelectric generator.

188

- Given the necessary materials and equipment, set up thermoelectric systems; and measure thermoelectric coefficients for both metallic conductors and semiconductors.

DISCUSSION

Thermoelectricity is the direct conversion of heat energy into electrical energy (or the reverse; that is, direct conversion of electrical energy into heat). Thermoelectric effects are observed in electrically-conducting materials through which electrical current flows. There are two broad classes of electrical conductors that show thermoelectric effects; they are metals and semiconductors. Thermoelectric effects in semiconductors are larger than in metals, so that semiconductors are used in practical thermoelectric power generators.

There are three related thermoelectric effects: the Seebeck effect, the Peltier effect, and the Thomson effect.

THE SEEBECK EFFECT

When two different conductors are joined in a loop, and when a temperature difference exists between the junctions at which the conductors are joined, an electromotive force (EMF) is produced. This EMF is called a "thermoelectric EMF," or a "Seebeck EMF." The situation is sketched in Figure 1. The two different conductors — for example, wires of two different metals — are joined at their two ends. When the junctions are at different temperatures, T_1 and T_2, as shown, a voltage V is produced between points A and B. A voltage-measuring device can be inserted between A and B. As long as the temperature at points A and B is the same, the thermoelectric voltage V will not be affected.

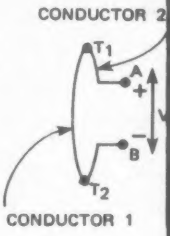


FIG. 1

The thermoelectric voltage, or EMF, is expressed by Equation 1.

$$V_s = (S_1 - S_2)\ (T_1 - T_2) \qquad (1)$$

where

V_s = Seebeck voltage, or net thermal EMF set up in a thermocouple.
S_1 = constant of conductor one.
S_2 = constant of conductor two.
T_1 = temperature at first junction.
T_2 = temperature at second junction.

The constant S for a material is called the "Seebeck coefficient" for the material. The Seebeck coefficient is dependent upon the temperature and may be either positive or negative. With reference to Figure 1, if $T_1 > T_2$ and the EMF has the indicated polarity, then $S_1 > S_2$. If the direction of the temperature gradient is reversed, the sign of the EMF also will be reversed.

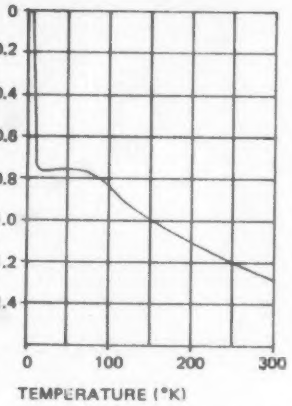


FIG. 2

According to Equation 1, if one measures the EMF and the temperature difference $(T_1 - T_2)$, one obtains $(S_1 - S_2)$, the difference in the Seebeck coefficients for the two materials. In order to obtain the individual coefficients, one makes measurements in which one of the materials is a reference material for which S is known. By convention, lead has been chosen as a standard reference material. The value of the Seebeck coefficient for lead is sketched in Figure 2 for temperatures from 0 to 300 K. The Seebeck coefficient varies rapidly at low temperatures, and becomes zero for temperatures below 10 K. The sign of the Seebeck coefficient for lead is negative.

Thus, to measure the Seebeck coefficient for a conductor, one may join the ends of the conductor to a reference conductor made of lead, as shown in Figure 1, and measure voltage and the temperature difference $(T_1 - T_2)$. Then the Seebeck coefficient for the unknown material can be found from Equation 1, by use of the known value for lead.

According to Equation 1, the voltage generated when two different materials are used in a thermoelectric circuit depends upon the difference $(S_1 - S_2)$, called the "thermoelectric power" for this combination of two different materials.

THE PELTIER EFFECT

The Peltier effect involves the heating and cooling of the junctions between two different conductors connected in a loop when an electric current flows in the loop. With reference to Figure 3, if the temperatures of the junctions are not permitted to change, the Peltier heat per unit time is released at the upper junction and absorbed at the lower junction. The Peltier coefficient π is defined by Equation 2.

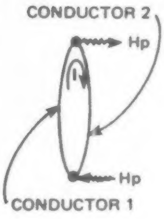


FIG. 3

$$H_p = (\pi_1 - \pi_2)I \tag{2}$$

where

H_p = Peltier heat released at upper junction and absorbed at lower.
π_1 = Peltier coefficient for conductor one.
π_2 = Peltier coefficient for conductor two.
I = current.

The effect is reversible; that is, if the direction of I were reversed, heat would be released at the lower junction and absorbed at the upper junction. For the direction of current flow as shown in Figure 3, with heat released at the upper junction, one would have $\pi_1 > \pi_2$.

The release of heat at one of the junctions leads to heating of the surroundings. The absorption of heat at one of the junctions leads to cooling of the surroundings. Thus, the Peltier effect can be used for either heating or for cooling.

THE THOMSON EFFECT

The Thomson effect is shown (Figure 4) as the release of heat from a single conductor in which a current is flowing, and in which a temperature gradient exists. The heat released is given by Equation 3.

$$H_T = \mu \, I \, (T_1 - T_2) \qquad (3)$$

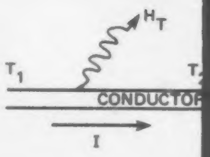


FIG. 4

where
H_T = Thomson heat.
μ = Thomson coefficient for conductor.

The effect is reversible; if the direction of current flow is reversed, then heat will be absorbed instead of released. The sign μ may be either positive or negative, depending upon the material. If $T_1 > T_2$, and if heat is released for the direction of current flow as shown in Figure 4, then $\mu > 0$.

The Thomson thermoelectric effect must be distinguished from ordinary resistive heating as given by I^2R, where R is the electrical resistance of the conductor. The I^2R heating is also present, of course, but its sign does not change as the direction of current flow changes. The Thomson thermoelectric effect is a different effect, which either can release or absorb heat in a conductor in a region of variable temperature.

These three thermoelectric effects are all related, and the coefficients for the Seebeck, Peltier, and Thomson effects are connected by two equations. The first of the equations is Equation 4, which can be stated as follows:

$$S = \frac{\pi}{T} \qquad (4)$$

Equation 4 relates the absolute temperature T and the Seebeck and Peltier coefficients. The second equation, which will not be presented here, is more complicated and depends upon a knowledge of differential calculus. This second equation relates the Seebeck and Thomson coefficients.

Thus, all three coefficients are related. If one coefficient is known, the other two may be determined by means of the two equations. For simplicity, the Seebeck coefficient usually is measured; then the Peltier and Thomson coefficients are derived from it. A simplified example involving a calculation of the Seebeck and Peltier effects is given in Example A.

EXAMPLE A: CALCULATION OF SEEBECK AND PELTIER COEFFICIENTS

Given: A thermocouple is set up as shown in Figure 1, with lead as conductor 2. $T_1 = 305$ K and $T_2 = 295$ K; and the observed voltage V, polarized as shown in Figure 1, is 30 µ volts.

Find:
(a) The Seebeck coefficient of conductor 1.
(b) The Peltier coefficient of lead at 300 K.

Solution: (a) From Equation 1:

$$(S_1 - [-1.28])(305 - 295) = 30\ \mu \text{ volts}$$

$$(S_1 + 1.28)(10) = 30\ \mu \text{ volts}$$

$$S_1 = 3 - 1.28$$

$$S_1 = 1.72\ \mu \text{ volt/K}$$

(b) Near 300 K, the Seebeck coefficient of lead is −1.28 µ volt/K, from Figure 2. From Equation 4:

$$\pi = S\ T$$

$$= (-1.28)(300)$$

$$\pi = -384\ \mu \text{ volt}$$

EFFICIENCY

The usefulness of a particular material for thermoelectric applications depends, not only upon its thermoelectric coefficient, of course, but also upon other material properties. For a good thermoelectric generator, the Seebeck coefficient should be large; but the electrical resistivity and thermal conductivity should be low. These competing demands lead to the concept of a "figure of merit" that combines all three properties and that expresses the utility of a material for thermoelectric applications. The figure of merit for a material is given by Equation 5.

$$Z = \frac{S^2}{\rho k} \tag{5}$$

where

Z = the figure of merit for a material.
S = Seebeck coefficient.
ρ = electrical resistivity.
k = thermal conductivity.

The figure of merit is useful in comparing different materials for thermoelectric applications. Some values will be given later in the description of materials. For semiconductors, Z may be in a range of about 0.001.

A thermoelectric generator, of course, utilizes two different materials. The expression for the figure of merit for the two materials used in combination is more complicated than that indicated in Equation 5; but it still combines the properties of Seebeck coefficient, electrical resistivity, and thermal conductivity for the two materials.

The figure of merit also is important because it affects the efficiency of a generator. The efficiency of a thermoelectric generator is the ratio of the electrical power to the rate of heat supply from the hot source. It expresses the fraction of heat energy converted to electrical energy. The maximum efficiency that can be obtained from a thermoelectric generator working between a heat source at temperature T_1 and a heat sink at a lower temperature T_2 is indicated by Equation 6.

$$\eta_{max} = \left(\frac{T_1 - T_2}{T_1}\right)\left(\frac{M - 1}{M + T_2/T_1}\right) \quad (6)$$

where

η_{max} = maximum efficiency obtainable from a thermoelectric generator.
$M = [1 + Z(T_1 + T_2)/2]^{1/2}$.

Some values of efficiency calculated with the aid of Equation 6 are shown in Figure 5 for several values of the figure of merit Z. The efficiency is plotted versus the temperature T_1 of the hot junction, for constant temperature T_2 = 300 K. For reasonable values of Z and T_1, efficiencies between 20% and 30% can be obtained.

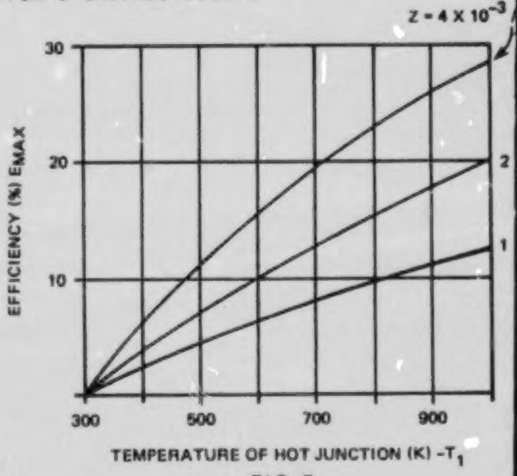
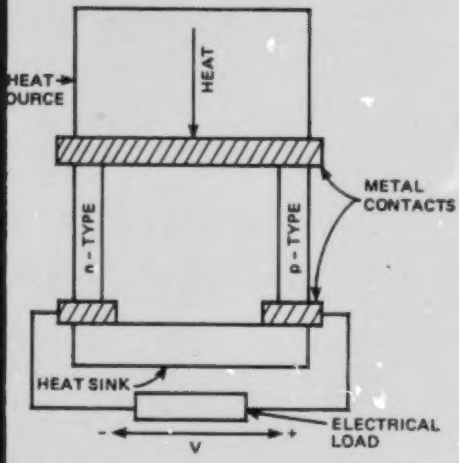


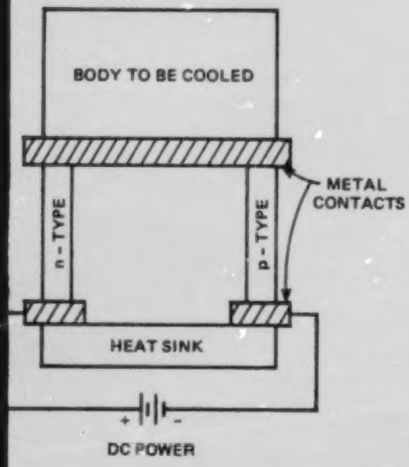
FIG. 5

The value of the efficiency given by Equation 6 is obtained only when the electrical resistance of the circuit is properly adjusted for optimum efficiency. The details for the proper adjustment are not presented here; however, the student should note that without such adjustment, the efficiency will be lower than the maximum value given by Equation 6.

CONSTRUCTION OF THERMOELECTRIC DEVICES



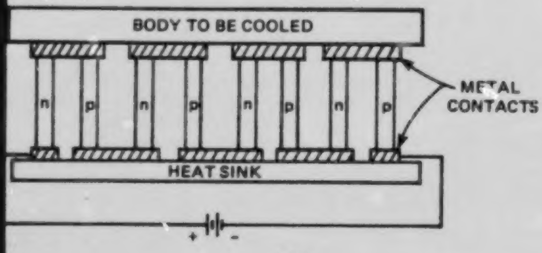
Now the construction of some thermoelectric generators shall be considered. Semiconductors exhibit larger thermoelectric effects and more complexity than metals. An electrical generator often is constructed of p-type and n-type semiconductor material (Figure 6). The Seebeck coefficient is positive for the p-type material, and negative for the n-type material. Thus, if S_2 corresponds to the n-type material, the term (S_1 - S_2) will be large in magnitude (because subtraction of the negative S_2 is the same as the addition of a positive value). Heat flows in from the heat source to the metal contact at the top, and is removed by the heat sink at the bottom. A temperature difference is established, and a voltage is generated in the external circuit, as indicated. This assembly, therefore, can generate electricity directly from the flow of thermal energy.



Thermoelectric effects also can be used for heating or for cooling, by virtue of the Peltier effect. Figure 7 illustrates a simple thermoelectric cooler, also utilizing semiconductors. With the current flow as shown, heat is withdrawn from the top junction, which is cooled. If the direction of heat flow were reversed, the top junction would be heated.

The basic elements displayed in Figures 6 and 7 often are combined in modules in which individual elements are connected in series electrically and in parallel thermally, as in Figure 8. This configuration allows development of coolers that have large areas, but that use only small currents.

When the required temperature difference for a Peltier cooler becomes large, stages frequently are cascaded (Figure 9). Then the hot junction of the first stage becomes the cold junction for the second stage. This arrangement allows the production of larger temperature differences than would be possible in a single-stage device.



THERMOELECTRIC MATERIALS

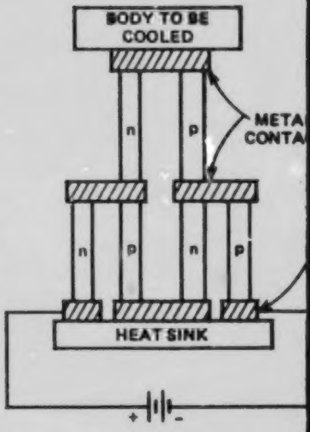


FIG. 9

Metals and semiconductors are the two broad classes of conducting materials used for thermoelectric applications.

METALS

Metals are convenient to use in wire form and, hence, are easier to fabricate into thermoelectric elements than are semiconductors. A thermocouple can be formed simply by soldering together the ends of two wires of different metals. The electrical leads are the wires themselves.

However, metals have relatively low values of Seebeck coefficient. Typical values lie in the range of microvolts per degree. These values are large enough to permit the use of metallic thermoelectric elements as thermocouples for temperature measurement, but are too small to allow metals to be efficient thermoelectric power generators.

There are a number of combinations of metals and alloys used in science and industry for temperature measurement. Table 1 lists some widely-used combinations. These combinations are chosen to have relatively large values of thermoelectric power; that is, of $S_1 - S_2$, as defined in Equation 1. The values of thermoelectric power range up to several tens of microvolts per degree (dependent upon temperature). Combinations of metals are available to cover various temperature ranges, as indicated in Table 1.

TABLE 1. METALS USEFUL FOR THERMOCOUPLES

Metal	Useful Temperature Range	Approximate Thermoelectric Power
Copper-Constantan[a]	-188°C to 350°C	15 - 60 μV/°C
Iron-Constantan[a]	200°C to 800°C	55 μV/°C
Platinum- Platinum + 10% Rhodium	0°C to 1500°C	6 - 12 μV/°C
Chromel-Alumel[b]	500°C to 1200°C	40 μV/°C

[a] Constantan is 55% copper and 45% nickel.

[b] Chromel is 90% nickel and 10% chromium; Alumel is 94% nickel, 3% manganese, 2% aluminum, and 1% silicon.

SEMICONDUCTORS

Semiconductors have much larger values of thermoelectric power than metals. Values range up to millivolts per degree, compared to a few microvolts per degree for metals. Thus, semiconductors are suitable for use in thermoelectric power generation and in thermoelectric coolers.

Semiconductors are not fabricated easily into wires and, thus, are less convenient to use than metals. Semiconductors are used in rod-or-bar form, and electrical lead wires must be attached.

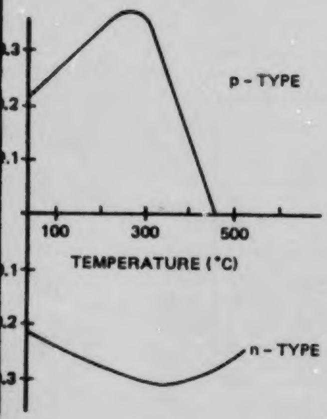


FIG. 10

One of the most widely-used semiconductor materials is lead telluride, which is a compound of the chemical elements lead and tellurium. Its chemical symbol is PbTe. The Seebeck coefficient for lead telluride is presented in Figure 10. The values in Figure 10 are for specific levels of impurities in the material; because the values of Seebeck coefficient are dependent upon the impurities in the material, these values must be considered as approximate. The Seebeck coefficient is positive for p-type material and negative for n-type material; therefore, a thermoelectric circuit can be formed from p-type and n-type PbTe. The thermoelectric power, $S_1 - S_2$, may be as large as 0.6 millivolts/degree at temperatures near 200°C.

There are a number of other semiconductor materials that can be used as thermoelectric elements. Some of these are listed in Table 2, along with representative values of the Seebeck coefficient. This table shows some of the materials that could be employed as thermoelectric generators.

TABLE 2. VALUES OF SEEBECK COEFFICIENT FOR SEMICONDUCTORS

Material	Seebeck Coefficient (mV/K)	Temperature (°C)
Zinc Antimonide (ZnSb) (p-type)	0.25	327
Bismuth Telluride (Bi_2Te_3) (p-type)	0.265	20
Indium Antimonide (InSb) (n-type)	-0.35	20
Indium Antimonide (InSb) (p-type)	0.80	-173

Lead Selenide (PbSe) (p-type)	0.26	27
Lead Selenide (PbSe) (n-type)	-0.29	27
Selenium doped with copper	1.24	--

Table 3 shows some values for the figure of merit Z (as defined by Equation 5) for several semiconductors.

TABLE 3. VALUES OF FIGURE OF MERIT

Material	Figure of Merit (per degree K)	Temperature (K)
PbTe (p-type)	0.0018	600
PbTe (n-type)	0.0019	500
ZnSb (p-type)	0.0012	500
Bi_2Te_3 (p-type)	0.0022	300
Bi_2Te_3 (n-type)	0.028	300

APPLICATION OF THERMOELECTRIC DEVICES

A variety of applications for thermoelectric effects have been developed.

TEMPERATURE MEASUREMENT

This is an old and well-established application in science and industry. Efficiency of energy conversion is not particularly important for this application. Thus, metal wires generally are employed in this application, rather than semiconductors, since wires are more convenient to use.

GENERATION OF ELECTRICITY

Thermoelectric generators of electricity offer the advantages of having no moving parts and of having high efficiency for small generators. Thermoelectric generators of less than a few kilowatts have efficiencies of up to 10%, a value higher than that available from conventional generators of comparable size. For larger generators, steam turbines have efficiencies of up to 40%, and thermoelectric generators do not compete well with them. Thus, thermoelectric generators are useful for generation of electrical power in remote areas, particularly in cases in which power requirements are not too high. Thermoelectric generators have been used as sources of electrical power in Arctic weather stations, navigation buoys, oil platforms in the ocean, and unattended lighthouses.

In the rural areas of the United States, electrification was completed some time ago, and thermoelectric sources did not achieve great importance. In rural areas of the Soviet Union, however, kerosene-fueled thermoelectric generators are an important source of electrical power.

SPACECRAFT POWER

Generation of electricity on spacecraft can be regarded as a special case. The heat source is a radioactive isotope. Such generators have been providing electrical power on orbiting spacecraft since 1961. They also have been used to power apparatus left on the moon by the Apollo astronauts.

COOLING

Small thermoelectric coolers have been developed for a number of applications. These coolers use the Peltier effect and produce cooling directly from the flow of electric current through the device. Small, single-stage devices can produce a temperature difference of 70C°, and two-stage devices can produce temperature differences greater than 100C°. Thermoelectric coolers have found many uses for scientific applications; for example, in the cooling of electronic components, such as transistors, lasers, and infrared detectors. They are employed in the cooling of specimens for biological research.

A few consumer products using thermoelectric coolers have reached the market; for example, small refrigerators, coolers on food-serving carts, and baby-bottle coolers. The high cost of thermoelectric coolers has kept them from being used for larger refrigerators.

HEATING

If the direction current flow in a thermoelectric cooler is reversed, the cooler becomes a heater. Thermoelectric devices have been developed for use as temperature-control devices; for example, housings for quartz-crystal oscillators have been controlled by use of thermoelectric devices to provide either heating or cooling as needed.

There have been a few consumer products, such as a "hostess cart," which provides both thermoelectric heating and cooling for different foods. Thermoelectric heating does not appear attractive for large-scale heating applications.

SUMMARY

Thermoelectricity is the direct conversion of heat energy into electrical energy (or the reverse; that is, direct conversion of electrical energy into heat). Thermoelectric effects are observed in electrically-conducting materials through which electrical current flows. There are two broad classes of electrical conductors that show thermoelectric effects; they are metals and semiconductors. Thermoelectric effects in semiconductors are larger than in metals, so that semiconductors are used in practical thermoelectric power generators.

There are three related thermoelectric effects: (1) The Seebeck effect (the reverse of the Peltier effect) indicates that when two dissimilar metals are heated in an electrical circuit, an EMF is produced and absorbed across the junction. (2) The Peltier effect, consists of EMF developed at the heated junction. (3) The Thomson effect is the EMF produced in a metal not heated uniformly.

THERMOELECTRIC VOLTAGE GENERATED BY A WIRE THERMOCOUPLE

MATERIALS

Potentiometer (for example, Leeds and Northrup, model K3); a digital multimeter (3 1/2 digit or better) or strip-chart recorder may be used in place of the potentiometer.
Thermocouple wires (for example, copper-constantan or platinum-platinum plus 10% rhodium).
Beakers.
Ice.
Water.
Bunsen burner and stand.
Electric wiring, clamps, supports, drill, bit for drilling metal, metal plates, solder, and soldering iron, as needed.
Thermocouple table that gives output of a particular thermocouple. (The Handbook of Chemistry and Physics contains such tables.)

PROCEDURE

1. Set up apparatus as illustrated in Figure 11. Use wires of common thermocouple materials; for example, copper and constantan. Twist together the ends of the wires, and solder them. A microvoltmeter may be used in place of the potentiometer, but a potentiometer will give more accurate results.

2. Measure the thermally-generated EMF.

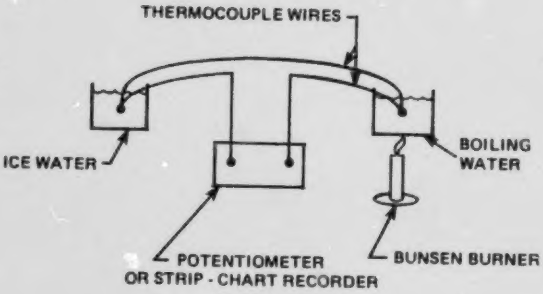


FIG. 11

3. Use a thermocouple table that gives the output of the particular thermocouple to compare the measured value to the tabulated value.

4. Use the results of Step 2 and Equation 1 to calculate the thermoelectric power, $S_1 - S_2$, for the pair of thermocouple wires.

5. Which one of the following effects is demonstrated by this experiment? (1) The Seebeck effect? (2) the Peltier effect? or (3) the Thomson effect?

THERMOELECTRIC VOLTAGE GENERATED BY SEMICONDUCTOR ELEMENTS

MATERIALS

Two bars of thermoelectric semiconductor, one p-type and one n-type (for example, Lead telluride from American Smelting and Refining Co.).
Potentiometer (for example, Leeds and Northrup, Model D3); a digital multimeter (3 1/2 digit or better) or strip-chart recorder may be used in place of the potentiometer.
Thermocouple from first experiment.
Hair dryer.
Metal plates.
Beaker.
Ice.
Water.
Electrical wiring.
Drill and bit for drilling metal.
Thermocouple table as in first experiment.

PROCEDURE

1. Set up apparatus as in Figure 12. Contact between the semiconductor bars and the metal plates is formed by firmly clamping the bars between the plates. The plates and bars must be solidly in contact. The thermocouple constructed in the first experiment is attached to upper plate. It can be embedded in a hole drilled in the plate.

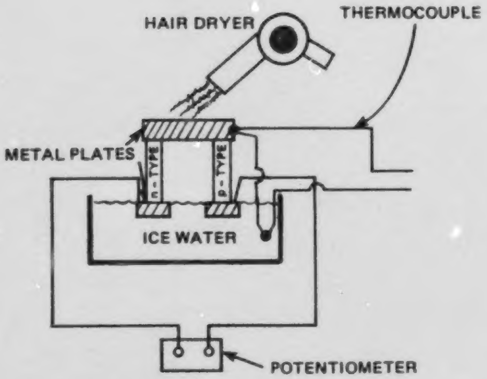


FIG. 12

2. Turn on hair dryer, and heat upper plate.

3. Measure thermoelectric voltage with potentiometer.

4. Disconnect semiconductor elements from the potentiometer, and connect thermocouple wires. Measure thermocouple voltage.

5. From Step 4 and calibration in the first experiment, determine temperature of upper plate.

6. Determine thermoelectric power, $S_1 - S_2$, for semiconductor elements. Use the results of Steps 3 and 5, along with Equation 1.

7. Compare the result from Step 6 with tabulated values. (Refer to the reference materials, or to Table 2 or Figure 10.)

COOLING PRODUCED BY A THERMOELECTRIC COOLER

MATERIALS

Thermoelectric cooler (available from Borg-Warner Thermoelectrics, Des Plaines, Illinois, and from Marlow Industries, Garland, Texas).
Two bars of thermoelectric material as in experiments one and two.
Thermocouple from first experiment.
Potentiometer, digital multimeter, or strip-chart recorder as in experiments one and two.
Beaker.
Ice.
Water.
Battery.
Drill and bit for drilling metal.
Metal plates.
Electrical wiring.

PROCEDURE

1. Set up apparatus as in Figure 13. The thermoelectric cooler is a commercial unit. Drill a small hole in the upper metal plate, and insert the thermocouple from first experiment. Hook up battery; make sure that voltage does not exceed maximum value specified by manufacturer.

2. Use potentiometer to determine voltage from thermocouple. Wait until top plate has cooled as much as possible.

3. Use calibration from first experiment to determine temperature of top plate.

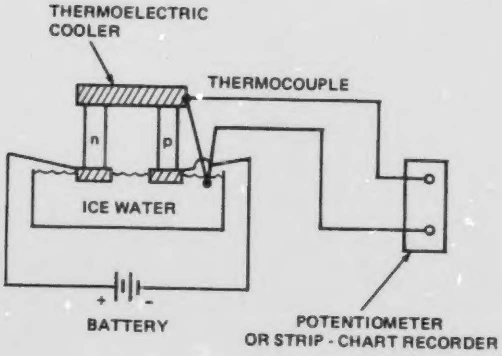


FIG. 13

4. Compare result to temperature drop claimed by manufacturer. Are there differences between the experimental results and the specified? If so, why?

SELF CHECK

1. Describe the Seebeck effect, the Peltier effect, and the Thomson effect.

2. Name two general types of thermoelectric materials. Name at least one specific example of each.

3. Name at least four applications of thermoelectricity.

4. The Seebeck coefficient of p-type PbTe is 0.35 mV/K, and of n-type PbTe is -0.30 mV/K. If these two materials are used in a thermoelectric circuit between a hot junction at 600 K and a cold junction at 350 K, how large is the thermoelectric voltage?

5. A thermoelectric generator, with figure of merit of 0.0032/K works between a hot junction at 950 K and a cold junction at 273 K. It has its electrical resistance adjusted for optimum efficiency. What is its efficiency?

6. The difference in temperature between two junctions is 60°C. By what factor does the Seebeck coefficient of one junction exceed the other if the Seebeck voltage is 400 mV?

7. What is the Peltier coefficient of a conductor at 16°C that has a Seebeck coefficient of 0.30 mV/K?

8. Find the figure of merit for a material that has a maximum efficiency of 22% and works between a hot junction at 800 K and a cold junction at 273 K.

SOLAR COLLECTORS

On a clear day, when the sun is directly overhead, solar energy arrives at the earth's surface at a rate of about 100 milliwatts/cm^2. Solar energy can be an important source of energy for space heating, water heating, cooling, industrial processing and electrical generation. This module discusses the solar collectors needed to harness this energy by converting energy from the sun into thermal or electrical energy. Solar collectors can take the form of simple, flat plates that absorb energy that heats water contained in pipes within the plates. The heated water then is used for space heating or for other applications. More advanced collectors include the concentrating time that focuses the sunlight, and the photovoltaic type.

In the laboratory, the student will construct a small section of a flat-plate collector, measure its efficiency, and also will measure the conversion efficiency of a photovoltaic collector.

OBJECTIVES

Upon completion of this module, the student will be able to:

- Discuss flat-plate collectors and concentrating collectors, including these factors:

  How they are constructed.
  At least three advantages for each type.
  At least three applications for each type.

- Define solar constant, give its numerical value, and state how much energy per unit area reaches the earth's surface on a clear day when the sun is directly overhead.

- Describe how photovoltaic solar collectors operate, and include a description of the photovoltaic effect.

- Given the necessary information, calculate the conversion efficiency of flat-plate and photovoltaic collectors.
- Given the appropriate equipment, measure the efficiency of a flat-plate collector and of a photovoltaic collector.

DISCUSSION

The sun delivers a large amount of energy to the earth in the form of electromagnetic radiation. This energy falls mostly in the visible, infrared and ultraviolet portions of the spectrum. It is, of course, responsible for the existence of life on earth. It makes plants grow. It provides much of the driving force for the circulation of the atmosphere and the oceans. Variations in the amount of solar energy reaching the earth lead to the progression of the seasons.

Energy delivered by the sun also can be employed for heating homes, offices and factories; for cooling; for water heating; and for industrial processes. The sun, thus, can be an important source of energy that can be harnessed to serve the needs of human beings. Good collectors for the incoming energy are needed. The discussion in this module concentrates on some of the more common collectors.

First, how much energy arrives from the sun? The amount reaching the vicinity of the earth (outside the earth's atmosphere) is called the "solar constant," which has a value of 135 milliwatts/cm^2. This number varies slightly as the earth revolves around the sun during the course of the year, but is a good average value. Thus, on a flat surface facing the sun and located above the atmosphere, each square centimeter (cm) would be insolated by 135 milliwatts. The word "insolation" shall be used to describe delivery of solar energy.

Some of the incident energy is lost during passage of light through the atmosphere. Some light is absorbed by molecules in the atmosphere; other light is scattered in such a manner that it changes direction. Some of the scattered light reaches the surface of the earth, but it will not be in the form of the direction beam from the sun.

Figure 1 shows, in a simplified way, the effects of the passage of sunlight through the atmosphere. Figure 1 plots the relative fraction of incident energy per unit area versus the wavelength. (Visible light covers the wave length range from 0.4 to 0.7 micrometers. Ultraviolet light has shorter wavelengths, and infrared radiation has longer wavelengths.) The reduction by the atmosphere of the energy

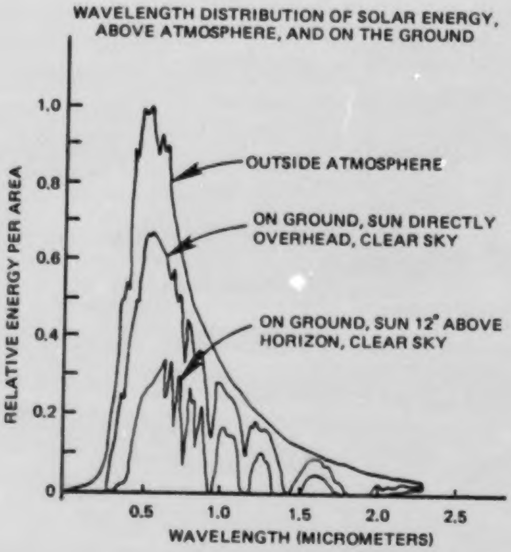

WAVELENGTH DISTRIBUTION OF SOLAR ENERGY, ABOVE ATMOSPHERE, AND ON THE GROUND

FIG. 1

reaching the surface is obvious. As the sun gets closer to the horizon, the energy must pass through a thicker layer of atmosphere, and less energy reaches the surface. If haze or clouds are present, the incident energy will be reduced even more than Figure 1 indicates. Figure 1 is also important for defining some of the properties of the collector. The solar collector must absorb at wavelengths matching those at which the solar energy reaches the ground.

The dips in the curves for energy reaching the ground as compared to energy outside the atmosphere are attributable to absorption by various molecules present in the atmosphere.

INCIDENT SOLAR ENERGY ON FLAT SURFACES AT GROUND LEVEL ON A CLEAR DAY

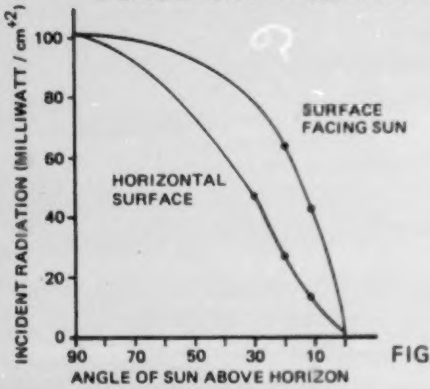


FIG. 2

Figure 2 shows the energy per unit area striking a flat area on the earth's surface on a clear day, versus the angle of the sun above the horizon. A surface tilted to face the sun can receive considerably more energy than a horizontal surface. Even in the best circumstance (that is, when the sun is directly overhead), the 135 milliwatts/cm^2 which reaches the top of the atmosphere is reduced to about 100 milliwatts/cm^2 by passage through the atmosphere. This value is equal to about 1 kW/m^2.

The total energy reaching the surface of the earth during the course of a year is impressive. Figure 3 shows the annual insolation over the United States. It is greatest in the Southwest; but even in the Northeast, it represents a large, potentially-useful energy source. If properly utilized, solar energy could reduce demands on other energy supplies.

ANNUAL ISOLATION IN THE U.S.
NUMBERS INDICATE KILOWATT HRS/m^2

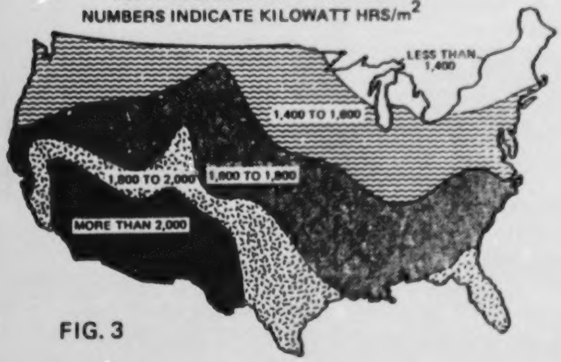


FIG. 3

FLAT-PLATE COLLECTORS

The most simple and basic collector is the flat-plate type, shown schematically in Figure 4. The absorbing surface collects the incident energy and thereby is heated. The tubes below the absorbing surface contain a fluid, usually water, that is circulated to deliver heat to an external source as needed. The transparent cover on the top allows incident light to reach the absorbing surface, but reduces reflective or radiative losses back to the atmosphere. The insulation on the back reduces heat loss by conduction. This simple, basic design has many variations,

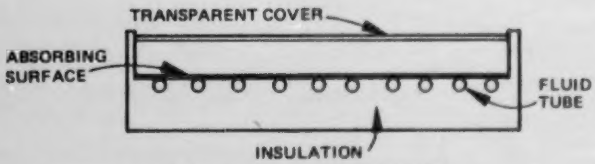


FIG. 4 BASIC FLAT PLATE COLLECTOR

but the more important features are illustrated in Figure 4.

The materials used in the flat-plate collector are important. Flat, black paint can be used for the absorbing surface, but efficiency can be increased by the use of selective absorbers, which absorb efficiently in the wavelength regions in which solar energy is concentrated. (See Figure 1.) Research is still underway to develop optimum materials for use in flat-plate collectors. With present technology, conversion efficiency of 50-60% from incident solar energy to thermal energy is possible with flat-plate collectors.

Figure 5 shows how a flat-plate collector could be used to provide space and water heating in a home. The collector is mounted on the roof, facing south and tilted to face the sun as much as possible. Pumps circulate the heated fluid as needed. The heat-storage reservoir could be a water tank. The heated water is circulated through the storage reservoir in pipes, and heats the water in the reservoir. This water, in turn, circulates through pipes in the house to provide space heating. Heated water direct from the collector is circulated through a water heater to heat water for home use.

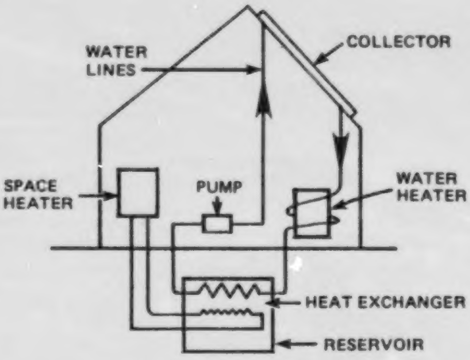


FIG. 5

There are several practical considerations that should be mentioned. First, a back-up conventional heating system probably will be needed in many areas, where a series of cold, cloudy days could make the solar heating system ineffective. Second, the water in the collector probably will cool to below freezing at night; therefore, a mixture of antifreeze and water normally is used. Finally, corrosion is a significant practical problem in working systems; therefore, corrosion inhibitors often are added to the water. The use of air as the heated fluid would eliminate the last two problems, but air is less efficient than water as a medium for heat exchange.

Solar energy also can be used to drive cooling systems. The cooling system is the same as that for a a gas refrigerator or air conditioner. The heated water takes the place of the flame in the gas refrigerator. Solar air conditioners of this type have been developed.

The temperature that can be reached with flat-plate collectors with current technology is around 150-200°C. This range is adequate for many practical applications. The applications that are now in use include the following:

- Space heating.
- Water heating.

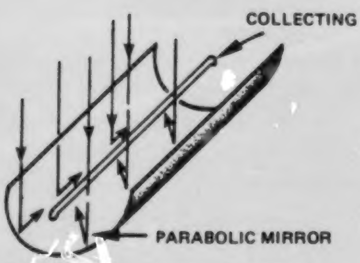
- Air conditioning.
- Industrial-process heat and industrial-process steam.

The cost of flat-plate solar collectors can be low enough to be competitive with conventional fuels. Reasonable estimates for the price of solar heating systems based upon flat-plate collectors are in a range of from $2 to $4 per square foot of collector. The National Aeronautics and Space Agency (NASA) has designed a system to heat a 1500-square foot house for about $2,040, based upon the dollar cost of materials in 1975. The estimate has been made that such an installation could pay for itself in ten years. Exact comparison becomes complicated and depends upon the climate in which it is used, costs for labor and materials, values for interest rates, payback rate for the initial investment and prices of conventional fuels in the future. For space heating, solar systems can be less costly than electrical heat or oil heat in the late 1970s, but probably still more expensive than heating with natural gas. As the price of natural gas rises in the future, however, solar heating probably will be less expensive than gas.

For water heating, solar heating is economically attractive in the late 1970s, compared to all three conventional fuels, because the system would be in use year-round. The efficienty of such systems coupled with a potentially rapid payback rate ensure a profusion of companies marketing solar water heating systems.

CONCENTRATING COLLECTORS

Because the temperature that can be reached with flat-plate collectors is limited, concentrating (or focusing) collectors sometimes are used. Such collectors utilize mirrors or other structures to focus solar radiation striking a large area onto an absorber of smaller area, which allows production of higher temperatures. One possible design of a concentrating collector is shown in Figure 6. A cylindrical reflecting mirror with a parabolic cross-section reflects the sunlight and focuses it onto a collecting pipe. The pipe is filled with a fluid (usually not water, because of the high temperatures reached). The fluid may be air or a molten salt. The pipe is surrounded by a quartz or glass envelope to reduce heat losses. The design shown in Figure 6 is one simple arrangement, but there are many other possible designs for concentrating collectors.



EXAMPLE OF FOCUSING COLLECTOR

FIG. 6

The temperature that can be reached in the collecting tube is much higher than that reached with flat-plate collectors. Temperatures up to 500°C are not difficult to produce within the pipe. In some research models, temperatures up to a few thousand degrees °C have been obtained. The efficiency of a concentrating collector can be higher than that for a flat-plate collector. It can range up to 75%. Thus, for a given, required power output, the size of the collector can be smaller if a concentrating collector is used.

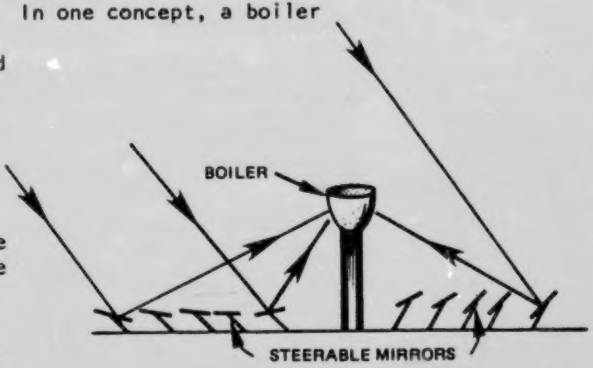
Concentrating collectors must track the sun, in order that the image of the sun is forcused accurately upon the absorber. Thus, they are mounted on movable mounts, and are equipped with solar sensors to move the collectors to follow the sun across the sky. Also, they depend upon a well-defined solar image, so that they almost are completely ineffective on a cloudy or a hazy day. Flat-plate collectors, on the other hand, will collect the diffuse radiation from the sky on a cloudy day, and will deliver some output.

Table 1 summarizes the relative advantages of flat-plate and concentrating collectors.

TABLE 1. RELATIVE ADVANTAGES OF FLAT PLATE AND CONCENTRATING COLLECTORS

Flat-Plate Collector	Concentrating Collector
Simpler construction	Higher efficiency
No need to track sun	Higher temperatures possible
Ability to use diffuse sky radiation	Smaller areas needed

Concentrating collectors can be employed for electrical power generation at a large, central station. In one concept, a boiler is mounted on a tower. A large number of mirrors covering a large area of land focus sunlight onto the boiler. The mirrors are steered independently by solar sensors. The concept is shown schematically in Figure 7. Solar power towers capable of driving turbines and generating megawatts of electrical power are under development. The mirrors may be spread over areas of more than a square mile.



TOWER CONCEPT FOR SOLAR POWER GENERATION

FIG. 7

In summary, possible applications of concentrating collectors include the following:

- Generation of industrial-process heat and steam at higher temperature than flat-plate collectors can provide.
- Cooking. Simple solar concentrators are in wide-spread use for cooking in many underdeveloped countries.
- Direct electrical-power generation.

PHOTOVOLTAIC COLLECTORS

The technology of solar collectors in practical use in the late 1970s is based upon the flat-plate and concentrating absorbers described above. The rapid development of photovoltaic collectors, however, offers the promise of direct generation of electricity from sunlight. Such systems may become practical and economical in the 1980s. The use of photovoltaic devices to generate electricity for spacecraft has been well-known for many years. However, such solar cells have been too costly for terrestrial applications. New manufacturing techniques now offer the possibility of economical production of large-area solar cells that could be used to generate electricity on the earth's surface.

As discussed in Application Module 9E5, "Photovoltaic Materials," the photovoltaic effect is produced as follows: At a junction between P-type and N-type semiconductor materials, there is an electric field in the semiconductor. When light is absorbed at the junction, it produces free electrical current carriers (electrons and holes) that move under the action of the electric field. This process produces an electric current. The photovoltaic effect allows direct conversion of the energy in sunlight to electrical energy. There is no need for an externally-applied voltage source.

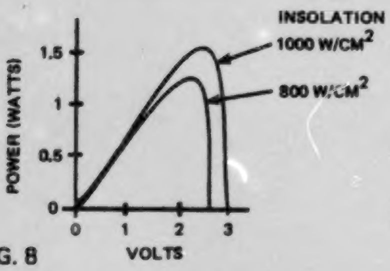


FIG. 8

Some characteristics of a silicon solar cell are shown in Figure 8, which shows the power output from a solar cell array of area around 20 cm^2 under two conditions of insolation. The maximum power is extracted at a particular level of voltage; at higher voltages the power drops off rapidly. The characteristics shown in Figure 8 are representative of what is available from state-of-the-art commercial devices.

The solar cells used on spacecraft were made of single-crystalline silicon. Silicon itself was inexpensive, since it is the second most common element in the earth's crust. However, processing costs were high , involving growth of a single crystal of silicon with high perfection. Now techniques are under development for making large areas of silicon solar cells inexpensively. The techniques include the dipping of a ceramic substrate into molten silicon to coat it with a thin layer. Other possibilities include the painting or spraying of silicon onto a substrate.

The key figure is the cost of the silicon solar cell. In order to make this technology economically feasible for terrestrial use, the cost of the solar cell must be reduced to about 50¢ per watt of electrical power. There is a large national program aimed at meeting this cost goal in the 1980s.

The low-cost solar cells described above could provide electrical generation for the home. A typical home might use about

600 kilowatthours of electricity each month. A photovoltaic system with a peak output of 5000 watts could meet this demand. In many parts of the country, a photovoltaic array with a working area of around 400 sq ft would be adequate. A battery storage system would be needed to carry over through the night.

Large area installations, perhaps in the southwestern United States, also would be possible for central power stations. Such systems covering hundreds of square miles in remote areas have been suggested. Such installations would be capable of generating a significant fraction of the electrical power requirements of the western United States.

The efficiency that has been demonstrated for large-area silicon solar cells is around 12-14%. Another possibility is gallium arsenide, which has a spectral response better matched to the solar spectrum shown in Figure 1. Gallium arsenide solar cells can achieve conversion efficiencies up of to 21%. They would be more expensive than silicon.

Solar photovoltaic systems offer a potential for providing significant amounts of electrical energy directly from sunlight. This conversion is clean and silent, and involves no heat engine and no moving parts. Its success will depend upon advances in manufacturing technology that will allow the production of solar cells at low cost.

SUMMARY

Solar collectors convert the energy of light radiation into thermal energy or into electrical energy. Flat-plate and concentrating (focused) collectors produce thermal energy in the fluid circulating through them. Photovoltaic collectors convert light radiation directly into electrical energy.

EFFICIENCY OF A FLAT-PLATE COLLECTOR

MATERIALS

Absorbing Material. (Foamglas absorber, 1½" slabs, sold by Pittsburgh Corning Corp., One Gateway Center, Pittsburgh, PA 15222, for 35¢/sq ft. Alternative: Roll-Bond Solar absorber. Sold by Olin Brass, East Alton, IL 62024, at $41.75 for 34 x 96" aluminum material, plus $25 packing fee for quantities under 50.)
Glazing material. (Kal-Lite fiberglass, 0.040" thick, from Kalwall Corp., Box 237, Manchester, NH 03105, at 40¢/sq ft.)
Wood (2" x 4" framing).
Duct tape.
Hammer, nails, saw and drill.
Flat-black spray paint.
Thermometer.
Pyranometer — desirable but not essential. Module 8-48 from the Eppley Laboratory, Newport, RI 02840, is suitable; cost $590.
Wristwatch.
Meter stick.

PROCEDURE

The purpose of this experiment is to construct a simple solar absorber, and to measure its efficiency.

1. Construct a simple solar absorber panel, according to the diagram in Figure 9. Many other types of absorbing materials and structures could be used but the design in Figure 9 is typical, and will illustrate the principles. The Foamglas absorber is suggested as a relatively inexpensive absorbing material. It also has the built-in insulator. Grooves help trap incident sunlight. The working fluid is air. Build a frame of 2 x 4 wood. Cut the Foamglas to size, and insert in the frame. Cut the glazing material to size, and attach to the top of the frame with duct tape. Drill a hole in the end of frame. Insert the thermometer through the hole in such a manner that its end is in air space above the absorber.

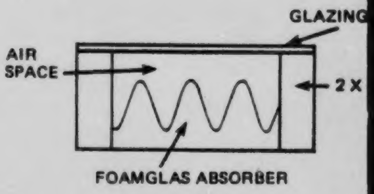


FIG. 9

2. The remainder of this experiment should be carried out on a clear, sunny day. Use pyranometer to measure the intensity

of the incident solar radiation. Record value in Data Table 1. If a pyranometer is not available, use Figure 2 to estimate intensity.

3. Move the absorber into sunlight. Place it horizontally on ground. In Data Table 1, record initial temperature reading of thermometer. After a few minutes, when temperature has risen perceptibly, record temperature again. Record elapsed time in seconds.

4. Move absorber inside, and allow it to cool to room temperature. Then move it back into the sun, and tilt it so that it faces the sun directly. Repeat step 3.

5. Move absorber inside. Allow to cool to room temperature. Remove duct tape and glazing, and spray paint absorber black. Allow paint to dry. Reinstall glazing and duct tape. Move absorber back into sun, and tilt so that it faces the sun directly. Repeat step 3.

6. Measure volume of the working fluid in the absorber. Volume V is approximately $V = \ell w\ (h_1 + \frac{1}{2} h_2)$ where ℓ is length of absorber, w is its width, h_1 is height from the glazing to top of absorber, and h_2 is depth of grooves in absorber. Measure these dimensions with meter-stick; record. Calculate V and record.

7. Fill out Data Table, and calculate efficiency η. As indicated in the Data Table, efficiency is defined simply as the ratio of energy absorbed by the heated air to energy delivered by the sun. The procedure, as outlined, allows measurement of absorbed energy as air is beginning to heat, when air is not far above ambient temperature. Measurement at this time minimizes losses due to conduction and reradiation. It allows measurement of transfer efficiency of solar energy.

8. How does the efficiency change as absorber is tilted from horizontal and pointed directly at the sun? How does it change as absorption is increased by black paint?

EFFICIENCY OF A PHOTOVOLTAIC COLLECTOR

MATERIALS

Photovoltaic silicon cell. (Model PIN-10DP [$49.00] from United Detector Technology, Santa Monica, CA, would be suitable.)
Voltmeter.
Ammeter.
Meter stick.
Operational Amplifier (Op Amp).

PROCEDURE

This experiment measures the efficiency of a silicon photovoltaic cell for sunlight.

1. Connect photovoltaic cell as indicated in Figure 10.

2. Point cell directly at sun. Measure current I and voltage V. Determine output power from this equation: P = IV.

3. Measure incident solar intensity with pyranometer. If a pyranometer is not available, estimate using Figure 2.

4. Measure area of photovoltaic cell (or use manufacturer's specification for area if available.).

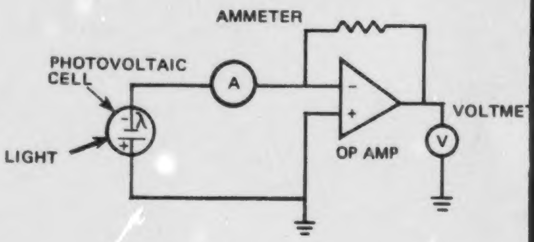


FIG. 10

5. Determine solar power incident on detector. (Area x result of step 3.)

6. Determine efficiency. (Step 2 divided by step 5.)

7. How does the efficiency of this photovoltaic collector compare with that of the flat-plate collector?

DATA TABLE 1: FLAT PLATE COLLECTOR

Incident Energy (Solar)

(a) Incident solar radiation ______ watts/cm^2 (from pyranometer)

(b) Area of solar panel ______ cm^2 (from measured ℓ & w)

(c) Time recorded for temperature change measured below ______ (See data below.)

(d) Incident energy = (a) ______ x (b) ______ x (c) ______ = ______ Joules.

Absorbed Energy (Working Fluid is Air)

	(a) Solar Panel Horizontal	(b) Solar Panel Tilted	(c) Solar Panel Tilted/Sprayed
Initial Temperature T_1	______ °C	______ °C	______ °C
Final Temperature T_2	______ °C	______ °C	______ °C
Elapsed time Δt between T_2 and T_1	______ s	______ s	______ s
Volume V of heated air $V = \ell w (h_1 + \frac{h_2}{2})$	______ cm^3	______ cm^3	______ cm^3
Mass m of heated air $(\rho_{air}V)(\rho_{air} = 0.00129$ gm/cm^3)	______ gm	______ gm	______ gm
Specific heat c_{air} of air $(\frac{cal}{gm \cdot C°})$ ($c_v$cair = 0.173 cal/gm·C°)	______ $\frac{cal}{gm \cdot C°}$	______ $\frac{cal}{gm \cdot C°}$	______ $\frac{cal}{gm \cdot C°}$
Absorbed Energy	$Q_a = MC\Delta T =$ ______ cal $Q_a =$ ______ cal $\times 4.18 \frac{J}{cal}$ $Q_a =$ ______ J	$Q_b = MC\Delta T =$ ______ cal $Q_b =$ ______ cal $\times 4.18 \frac{J}{cal}$ $Q_b =$ ______ J	$Q_c = MC\Delta T$ ______ cal $Q_c =$ ______ cal $\times 4.18 \frac{J}{cal}$ $Q_c =$ ______ J
Efficiency	$\eta \equiv \frac{Q_a}{E_{incident}}$ = ______ (See d above.)	$\eta \equiv \frac{Q_b}{E_{incident}}$ = ______	$\eta \equiv \frac{Q_c}{E_{incident}}$ = ______

DATA TABLE 2: PHOTOVOLTAIC COLLECTOR						
Current I	Voltage V	Power P	Incident Solar Intensity	Area of Cell A	Solar Power Incident on Detector	Efficiency η

SELF CHECK

1. Discuss flat-plate collectors and concentrating collectors. Include a description of how each is constructed, at least three advantages for each type and at least three applications for each type.

2. Define the solar constant. State its numerical value. State how much solar energy per unit area reaches the earth's surface on a clear day when the sun is directly overhead.

3. Describe how photovoltaic solar collectors operate; include a description of the photovoltaic effect.

4. A flat-plate solar collector has an area of 10 square meters; and it is pointed directly at the sun, which is 45° above the horizon. The collector, with its associated equipment, generates an output power of 5 kilowatts. What is its efficiency?

   HINT: Use Figure 2.

5. A photovoltaic collector has an area of 120 cm^2. When 100 milliwatts/cm^2 of solar radiation is incident on it, it generates 0.7 amperes of electrical current at a voltage of 2 volts. What is the efficiency of the collector?

6. At approximately what angle above the horizon is the sun if, when a 30 m^2 solar collector is pointed directly at the sun, 12 kW of power are generated? Assume the efficiency of the solar collector is 47%.

7. How large should a solar collector be if it is to produce 10 kW of power when 1 kW/m^2 is incident upon it and the collector is 10% efficient?

8. How much current can be produced by a solar collector which is 5% efficient and takes in 100 kJ of energy each second if its output voltage is 110 volts?

TRANSDUCERS

APPLICATION MODULES

		Page
10M1	Strain Gages	213
10M2	Accelerometers	225
10F1	Bourdon Tubes	232
10F3	Differential Pressure Transducers	241
10F4	Fluid-Flow Measuring Devices	250
10F5	Vacuum Gages	258
10E1	Meter Movements	275
10T2	Bimetallic Strips	288
10T3	Temperature Measuring Devices	299
10T4	Thermocouples	312
10T5	Thermistors	328
10T6	Pyrometers	340

10M1

STRAIN GAGES

The strain gage is a transducer device used in the measurement of small dimensional changes in a surface. It consists of a fine wire winding (or of foil or semiconductor materials) bonded firmly to the surface tested. As the surface deforms under strain, so does the strain gage, causing the dimensions of the wire to undergo change and, therefore, to change its electrical resistance. This change in resistance is measured to determine the strain experienced by the sample.

This module introduces the construction and operation of various strain gages. In the laboratory, the student will measure the change in resistance of a wire subjected to strain, and will use a strain gage to measure a sample strain.

OBJECTIVES

Upon completion of this module, the student should be able to:

- Briefly describe operation of a strain gage.
- Define the term "gage factor" and state ideal and typical real values for the gage factor.
- List the three types of materials used for strain gages and at least one specific metal alloy used.
- Calculate the gage factor of a wire, given the wire resistance, strain, and increase in resistance.
- Given the appropriate equipment, measure the change in resistance of a wire under tensile stress, and measure strain in a metal bar to which force is applied.

DISCUSSION

A strain gage is a device used to measure dimensional changes in a sample of material. The strain gage consists of a material whose electrical properties change as it is stretched or deformed. For example, a length of fine wire under tension will undergo a slight dimensional change accompanied by a change in electrical resistance of the wire. If the wire is bonded firmly to the sample, measurement of change in electrical resistance of the wire can be used to determine strain in the sample.

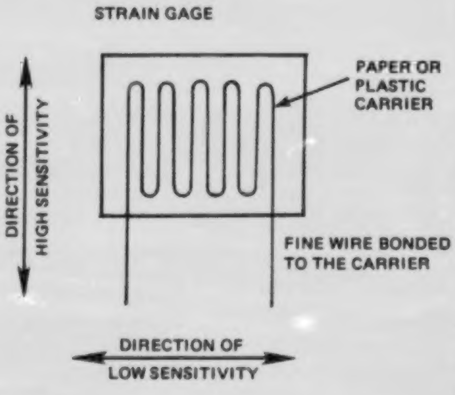

STRAIN GAGE
PAPER OR PLASTIC CARRIER
DIRECTION OF HIGH SENSITIVITY
FINE WIRE BONDED TO THE CARRIER
DIRECTION OF LOW SENSITIVITY

FIG. 1

TYPES OF STRAIN GAGES

Strain gages take a number of forms: thin wires, foils, or semiconductors. A typical form for a wire strain gage is shown in Figure 1. The fine wire is first bonded to a carrier material (e.g., paper or plastic) by cement. One particular shape for the wire is shown in Figure 1, but configurations in a wide variety are used in modern strain-sensitive elements. The gage then can be cemented to the sample. When the sample is strained, strain is transmitted to the strain gage. Measurement of electrical resistance of the wire gives the strain in the wire and, therefore, the strain in the sample to which it is bonded.

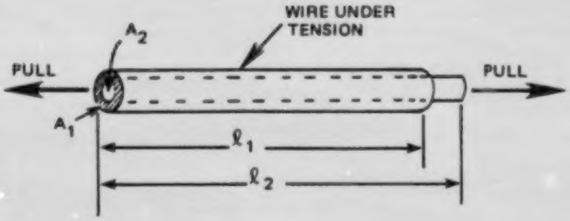

WIRE UNDER TENSION
A_2
A_1
PULL
PULL
ℓ_1
ℓ_2

FIG. 2

Figure 2 illustrates how the strain of a wire affects its electrical resistance, and defines length and cross-sectional area of a wire before it is strained. When the wire is strained, it becomes slightly longer, and the cross-sectional area becomes smaller. Electrical resistance before the wire is strained is given by Equation 1.

$$R_1 = \frac{\rho \ell_1}{A_1} \qquad (1)$$

where

R_1 = electrical resistance of wire before being strained.
ρ = electrical resistivity of wire.
ℓ_1 = length of wire before being strained.
A_1 = cross-sectional area of wire before being strained.

If the assumption is made for the moment that electrical resistivity is unaffected by strain, resistance after strain is applied is given by Equation 2.

$$R_2 = \frac{\rho \ell_2}{A_2} \tag{2}$$

where

R_2 = electrical resistance of wire after being strained.
ℓ_2 = length of wire after being strained.
A_2 = cross-sectional area of wire after being strained.

Thus, resistance changes because of dimensional changes in the wire. The wire becomes longer as it is strained, and the cross-sectional area decreases; both of these changes affect the electrical resistance, according to Equations 1 and 2.

Change in electrical resistance is related to strain, as given in Equation 3.

$$\frac{\Delta R}{R} = \frac{\Delta \ell}{\ell}(1 + 2\nu) = S(1 + 2\nu) \tag{3}$$

where

ΔR = change in original resistance R due to a change $\Delta \ell$ in original length ℓ.
ν = an elastic constant called Poisson's ratio. (Refer to next paragraph.)
S = strain, equal to $\Delta \ell / \ell$.

Poisson's ratio ν varies slightly in value from one metal to another. For most metals, it is approximately equal to 0.3, the value used in this module.

Since $\Delta \ell / \ell$ is recognized as strain S — a dimensionless quantity given sometimes as "inches per inch" or "centimeters per centimeter" — Equation 3 can be rewritten as shown in Equation 4 where the gage factor G is introduced.

$$\frac{\Delta R}{R} = GS \tag{4}$$

where

G = gage factor, $1 + 2\nu$.

Equation 4 is used to solve a problem in Example A.

Definition of the gage factor G follows from Equation 4. G is equal to fractional change in resistance $\Delta R/R$ divided by strain $S = \Delta \ell/\ell$; thus, the gage factor is a fractional change in resistance per unit strain. Since G is equal to $(1 + 2\nu)$, and ν for most metals is equal to 0.3, ideally the gage factor should equal 1.6. However, Equations 3 and 4 were derived on the assumption that electrical resistivity $\rho = RA/\ell$ does not change with strain. This is not quite true. Electrical resistivity ρ does change with strain in a rather complicated manner. When this change is taken into account, the gage factor G is typically found to have values in the range from $G = 1.8$ to $G = 4.5$ rather than the simple value of 1.6 given above.

To summarize, the relationship between fractional change in resistance and strain S is given by Equation 4. Ideally, G should have the value 1.6, but in practice the value of G will vary between 1.8 and 4.5 because of the change of electrical resistivity with strain (see Example A).

EXAMPLE A: RESISTANCE CHANGE CALCULATION

Given: A strain gage has a resistance of 1000 Ω; gage material has a value of modulus of elasticity of 10^{12} dyne/cm^2 and is subjected to a stress of 10^9 dyne/cm^2, and a gage factor (G) of 2.0 is assumed.

Find: Change in resistance.

Solution: Strain S is found from Hooke's Law equation*:

$$E = \frac{P}{S}$$

$$S = \frac{E}{P}$$

where
P = stress.
E = modulus of elasticity.

Using values above,

$$S = \frac{10^9}{10^{12}}$$

$$S = 10^{-3}$$

Then, using Equation 4, change in resistance ΔR is derived as follows:

Thus,

$$\frac{\Delta R}{R} = GS$$

$$\Delta R = GSR$$

$$\Delta R = 2 \times 10^{-3}\ (1000\ \Omega) = 2\Omega$$

---

*Hooke's Law is covered in detail in Module IM5.

MATERIALS AND CONSTRUCTION OF STRAIN GAGES

Strain gages can be made of three different types of materials:

Small-diameter wire (about 0.001 inch diameter).
Thin sheets of metal foil (about 0.00015 inch thick) etched to form patterns similar to that in Figure 1.
Semiconductor materials, basically similar to metallic gages, but giving greater response for a given strain.

The arrangement for strain-gage design shown in Figure 1 gives relatively high sensitivity to strain along the length of the loops of wire and low sensitivity to strain perpendicular to that direction. For two-dimensional measurements, two separate, electrically insulated, gages, oriented in two perpendicular directions, can be used.

Commercial strain gages come with the sensitive element mounted on a carrier or backing, which serves both to support the unit and to provide electrical insulation from the test specimen. The gage is bonded rigidly to the specimen by an adhesive, in such a manner as to transmit the strain from the specimen to the gage (Figure 3). In this configuration, the device commonly is called a "bonded" strain gage. An unbonded strain gage could involve wires stretched between supports on two components that move relative to each other. The bonded type of strain gage has become more common.

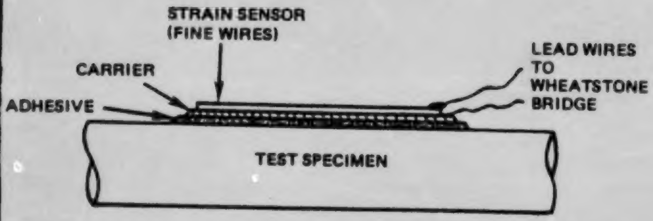


FIG. 3

A few of the metals used in metallic wire and foil gages are presented in Table 1, along with comments on useful temperature range and the recommended types of use. Situations where measurements are made of forces that are of a steady state or the rate of change is low are labeled static measurements. Situations where the forces being measured are constantly or rapidly changing magnitude are labeled dynamic measurements.

TABLE 1. METALLIC MATERIALS FOR STRAIN GAGES

Metal	Temperature Range	Recommended Type of Measurement
Copper-nickel alloy (Constantan)	-100 to +460°F	Static
Nickel-chrome alloy (Nichrome V)	to 1200°F to 1800°F	Static Dynamic
Nickel-iron alloy (dynaloy)	Limited by carrier and adhesive	Dynamic
Nickel-chrome alloy (stabiloy)	to 600°F	Static
Platinum alloy (alloy 1200)	to 1200°F to 1500°F	Static Dynamic

For practical measurement of the relatively small resistance changes caused by strain, a Wheatstone bridge often is used, as illustrated in Figure 4. Unknown resistance of the gage is R_1; resistances R_2, R_3 and R_4 are known, with R_3 being a variable resistance. To find the value of the unknown resistance R_1, the variable resistor R_3 is adjusted until the galvanometer G registers no current flow; that is, $I_1 = I_3$, and $I_2 = I_4$. Under this balance condition, the four resistances must satisfy Equation 5.

$$R_1 = \left(\frac{R_2}{R_4}\right) R_3 \qquad (5)$$

R_1 can be determined by Equation 5 immediately from the known values R_2, R_3, and R_4. Commercially-packaged bridge networks are available; variable resistance may be adjusted simply by the rotation of dials until the meter on the instrument indicates zero current flow.

In practice, the bridge circuit may be set up as shown in Figure 5. A dummy gage is mounted on an unstrained sample of the same material that is being measured. The active gage is mounted on the strained material. This configuration compensates for changes in temperature, which affect both the active gage and the dummy gage identically.

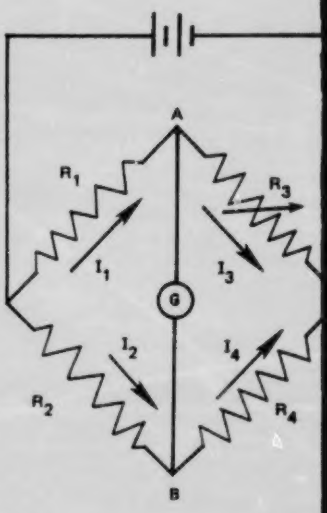


FIG. 4

Strain gages often are used in industry for measuring dimensional changes in specimens, as they offer the advantages of an easily-measurable electrical output, ease of installation and operation, and low cost. In addition to their direct use for measuring dimensional changes, they are utilized to determine the cause of strain.

SUMMARY

The strain gage is a transducer device used in the measurement of dimensional stress changes in a surface. The gage component (wire, foil, or semiconductor material) changes proportionally in resistance as strain changes upon the surface to which the gage is attached. This resistance is measured to determine the strain on the surface.

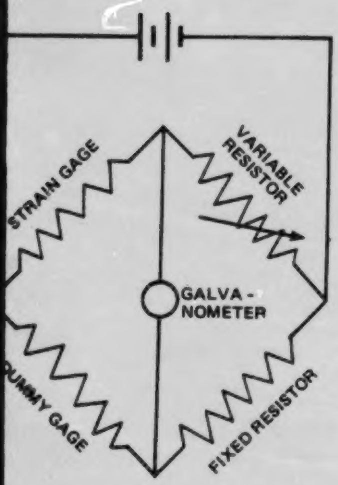


FIG. 5

NICKEL WIRE RESISTANCE CHANGE

MATERIALS

Nickel wire, diameter 0.5 mm x length 36 in.
Spring scale.
Wheatstone bridge. (Refer to Step 1.)
Heavy immovable object for mounting.
Micrometer, 0-1 in.

PROCEDURE

In this experiment, the student will measure change in resistance of a nickel wire while it is being subjected to a tensile stress.

1. Set up apparatus as shown in Figure 6. Procedures assume that a commercial Wheatstone bridge is used, but an improvised bridge can be contructed with known resistors and a galvanometer, as shown in Figure 4. Attach one end of nickel wire to a heavy immovable object. Attach spring scale to other end. Measure initial unstrained resistance of wire with bridge; enter in Data Table 1.
2. Apply a force with spring scale; measure resistance again. Enter values in Data Table 1.
3. Repeat step 2 for several different values of applied force on spring scale, entering values in Data Table 1.
4. Measure diameter of wire with micrometer; calculate cross-sectional area. Check to see whether or not changes in radius with load are large enough to be taken into account.
5. For each value of applied force, calculate strain. (Remember that strain equals force per unit area divided by modulus of elasticity; for nickel, the modulus of elasticity is 3×10^7 pounds per square inch.) Enter values in Data Table 1.

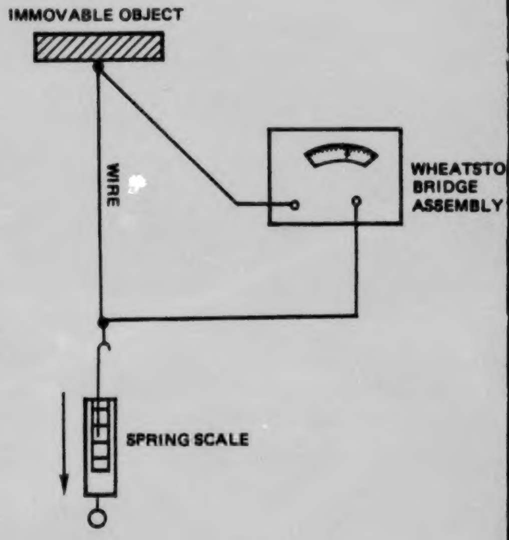


FIG. 6

6. What would be the gage factor of this wire if it were used as a strain gage?

METAL BAR STRAIN

MATERIALS

Spring scale.
Wheatstone bridge (as in previous experiment, step 1).
Metal bar (size-round 1/2 in diameter; square 1/4 in x 1/4 in x 48 in long).
Commercial bonded strain gage (model A-15X from BLH Electronics or equal).
Adhesive (Eastman 910 or equal).
Clamp.
Electrical lead wires and connectors, as needed.
Transparent pressure-sensitive tape, as needed.
Tissue, as needed.
Soldering iron and resin-core solder, as needed.

PROCEDURE

In this experiment, the student will use a commercial strain gage to measure strain in a metal bar.

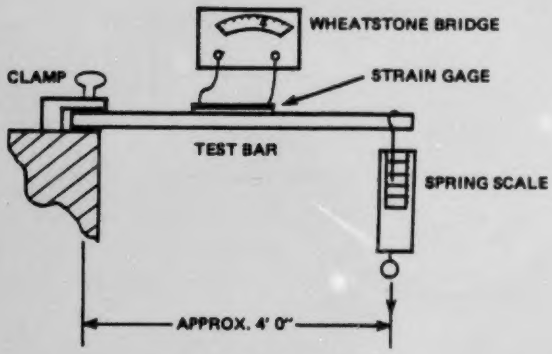


FIG. 7

1. Set up apparatus as shown in Figure 7. Clamp one end of metal bar rigidly with a clamp mounted on a heavy immovable object.

2. Cement strain gage to top of bar by following instructions provided by manufacturer. If no instructions are supplied, use following procedures:

   Pick up gage with a piece of pressure-sensitive transparent tape. Spread a thin layer of adhesive on back of gage. Align gage accurately over area to which it will be attached; remember that once gage touches surface, to move it again probably will not be possible. Apply one end of tape to within about 1/8 in of gage end. Then wipe down tape and gage with a steady, progressive pressure, smoothing down with tissue so that cement does not touch fingers. Wait one minute; peel off tape slowly. Attach lead wires by soldering to lead-wire pads, if lead wires were not attached originally.

3. Connect Wheatstone bridge and attach spring scale.
4. Measure initial resisance with zero force on scale.
5. Measure resistance with a force on the scale. Calculate change in resistance. Enter values in Data Table 2.
6. Repeat step 5 for several increments of force on the scale; enter values in Data Table 2.
7. For each value of resistance change, calculate strain. Use manufacturer's stated value of gage factor.
8. How does the strain gage used in this experiment operate?

DATA TABLE 1

Scale Reading (Pounds)	Strain	Resistance (ohms)

DATA TABLE 2		
Scale Reading (Pounds)	Change in Resistance (Ohms)	Strain

SELF CHECK

1. Briefly describe strain gage operation.

2. Define gage factor for a strain gage. State ideal value of gage factor, and list at least one typical real value.

3. List three types of materials used in strain gages. Name at least one specific metal used.

4. A strain gage has an initial resistance of 400 Ω. Under a strain of 10^{-3}, resistance increases to 401 Ω. What is gage factor?

5. The resistance of a wire is 10Ω when the wire has a length of 2 m and a cross-sectional area of 0.003 m^2. What is the resistance of the wire when it is stretched to a length of 3 m and has a new cross-sectional area of 0.002 m^2? Assume that the electrical resistivity is unaffected by strain.

6. Find the fractional change in resistance (ΔR/R) for a material subjected to a strain of 10^{-2}. (Take Poisson's ratio to be equal to 0.3.)

7. What is the resistance of a wire for which G = 2, if the wire had a resistance of 5Ω before being subjected to a stress of 10^{-4}?

8. A wheatstone bridge like the one shown in Figure 4 uses resistors such that; R_2 = 100Ω, and R_4 = 50Ω. Find R_1 if the galvanometer registers no current flow when R_3 = 200Ω.

10M2

ACCELEROMETERS

The piezoelectric accelerometer contains certain crystals that exhibit surface charges when subjected to mechanical stress. The magnitude of the charge is proportional to the stress. The accelerometer changes this mechanical stress energy into the electrical charge or voltage measured to determine acceleration.

This module investigates the construction and operation of the piezoelectric accelerometer. In the laboratory, the student will use an accelerometer to measure a variety of accelerations.

OBJECTIVES

Upon completion of this module, the student should be able to:

- Sketch and label the five major components of a single-ended compression accelerometer--including the critical axis--and describe in a sentence or two how it works.
- Sketch and label the three major components of a shear design accelerometer, its critical axis, and give a one- or two-sentence description of how it works.
- Sketch the voltage equivalent circuit of an accelerometer, and label the signal source, capacitance, and resistance.
- Name a type of accelerometer measuring device, and describe in one or two sentences its key characteristics.
- Given two of these three values, solve for the unknown quantity: acceleration, peak voltage, and sensitivity.

DISCUSSION

Figure 1 shows the components of a straight-compression and of a single-ended compression accelerometer. In each case, K is a piezoelectric crystal, and M is a mass held in contact with the crystal. A compression is maintained on the crystal at all times to increase the sensitivity of the accelerometer. When the device experiences an acceleration along the axis shown in the the figure, the pressure exerted on the crystal K by mass M changes. This pressure, in turn, changes the magnitude of the surface charge on the crystal. Electrodes on the crystal surface are connected to an electrical output to provide a means of measuring either the charge or the voltage produced.

While the straight-compression design is the simplest, it is seldom used in modern accelerometers. The single-ended compression design is the most common. It provides both electrical and mechanical isolation of the accelerometer from the case that encloses it. The cap of this design is a metal can welded to the accelerometer base; it seals the unit, but does not come into direct contact with the working parts of the accelerometer.

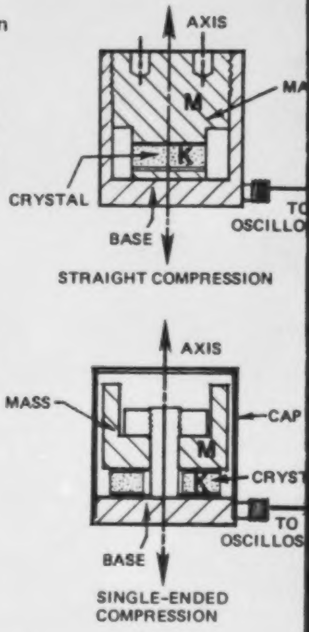


FIG. 1

Figure 2 shows a shear-design accelerometer. The crystal is a cylinder bonded to a hollow shaft. A concentric cylindrical mass is bonded to the crystal, and the entire assembly is enclosed. An acceleration along the axis shown in the figure results in a shear force within the crystal. The electrodes are located on the inner and outer walls of the cylindrical crystal, and are connected to an electrical output.

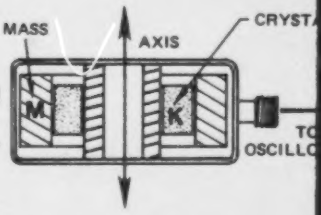


FIG. 2

Shear accelerometers can be built with extremely small dimensions and can be mounted by the passing of a machine screw through the central hole. Most microminiature accelerometers are of this design.

Measurement of the charge collected on the electrodes of the crystal is the most accurate measurement of acceleration. Because such measurement is both difficult and expensive, however, voltage measurements usually are made. Figure 3 shows the voltage equivalent circuit of an accelerometer and its associated voltage-measuring device. C_p is the capacitance of the accelerometer electrodes, and C_t is the capacitance of the cable used to connect the accelerometer to the measuring device. R_L is the input impedance of the voltage-measuring device (a vacuum-tube voltmeter, called a VTVM, or an oscilloscope). The charge generated by the crystal is q. For the best response, $C_p + C_t$ should be as small as possible, and R_L should be as large as possible.

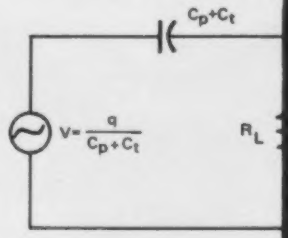


FIG. 3

Accelerometers usually are supplied with low-capacitance cables of known capacitance. The voltage sensitivity supplied by the manufacturer is based upon the use of these cables. The accelerometer will not be calibrated correctly if other cables are used.

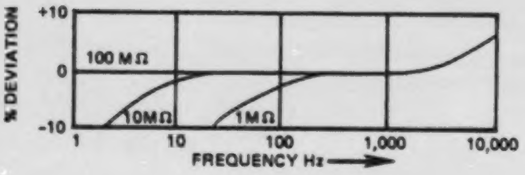


FIG. 4

Figure 4 shows the typical percentage of deviation from the calibrated voltage sensitivity of an accelerometer for three values of load resistance. This figure indicates that good low-frequency performance requires the use of a voltage-measuring device having a very high input impedance. Thus, an oscilloscope is the most common measurement device used.

Equation 1 is used to determine the acceleration measured by an accelerometer and oscilloscope combination.

$$a = \frac{V}{s} \qquad (1)$$

where

a = acceleration.
V = peak voltage in mV.
g = acceleration due to gravity.
s = sensitivity in $\frac{mV}{g}$.

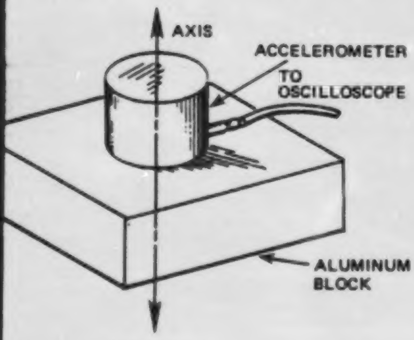


FIG. 5

The sensitivity usually is given in millivolts per g with a specified external capacitance (that of the connecting cable provided), and the resulting acceleration is measured in number of g's. Example A shows the use of Equation 1 in solving a problem.

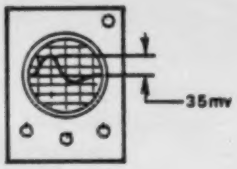
In practical applications, accelerometers must be firmly attached to the component whose acceleration is to be measured. If the accelerometer cannot be bolted directly to the component, it may be secured to a metal block held in contact with the component (Figure 5).

Recall from Figures 1 and 2 that accelerometers measure accelerations along one axis only. When mounted as shown in Figure 5, the accelerometer will respond only to those acceleration components that lie along its axis. Accelerations perpendicular to this axis cannot be detected by the device; therefore, the accelerometer can be used to measure both the magnitude and direction of accelerations.

EXAMPLE A: CALCULATION OF ACCELERATION USING ACCELERATION READINGS

Given: An accelerometer with a sensitivity of 16 mV/g was used to produce the oscilloscope trace shown in the drawing below.

Find: The acceleration.



Solution:

$$a = \frac{V}{s}$$

$$= 35 \text{ mV} \div 16 \frac{\text{mV}}{\text{g}}$$

$$a = 2.19 \text{ g}$$

The acceleration was 2.19 times the acceleration of gravity or $2.19 \times 32.2 \text{ ft/s}^2 = 70.5 \text{ ft/s}^2$.

SUMMARY

The piezoelectric accelerometer is a transducer that operates as follows: (1) Mechanical pressure, or stress, is applied to the accelerometer crystals during acceleration. (2) This stress energy is changed by the crystals into an electrical charge proportional to the stress. (3) The electrical charge or voltage is measured to obtain the amount of acceleration. Accelerometers can measure the magnitude and the direction of accelerations along certain axes.

ACCELEROMETER MEASUREMENTS

MATERIALS

Accelerometer (Endevco model 2215E or equal).
Oscilloscope with 0.005 V/cm sensitivity.
Aluminum block approximately 2" x 2" x 1" drilled and tapped to accept accelerometer stud.
Centimeter scale.
Several sources of mechanical vibration, such as drill press, electric shaver, electric fan, saw, etc.

PROCEDURE

1. Attach accelerometer securely to aluminum block, as shown in Figure 5. Use cable provided to connect accelerometer to input of oscilloscope.

2. Turn on oscilloscope, and obtain a stable trace on the scope display. Voltage sensitivity of the oscilloscope should be set for 0.005 V/cm. Note that this setting and time base of oscilloscope must be changed to the appropriate settings throughout the experiment to give the best display. Settings necessary will vary with sources of vibration used.

3. Place mounted accelerometer on a table, and tap block lightly with one finger while observing oscilloscope. Display shows variations in acceleration experienced by accelerometer. Tap block lightly with a pencil or screwdriver, and observe resulting vibrations.

4. Hold accelerometer 2 cm above table top, with its axis vertical. Drop accelerometer while observing oscilloscope. Record in Data Table amplitude and frequency of resulting accelerations.

5. Repeat above procedure with accelerometer axis horizontally.

6. Place accelerometer on floor, with its axis vertical. Have another student jump as high as possible and land a few feet from accelerometer. Record results from oscilloscope.

7. Repeat step 5 with accelerometer axis horizontal.

8. Place accelerometer mounting block in contact with a source

of mechanical vibration. Observe and record results.

9. Use same source of mechanical vibration, but rotate accelerometer axis by 90°. Record results.

10. Repeat steps 7 and 8 for another source of mechanical vibration.

11. Complete the Data Table.

12. If the oscilloscope did not display a voltage during some of the tests, or if it displayed a very small voltage, why?

DATA TABLE

Source of Acceleration	Accelerometer Axis	Peak Voltage (mV)	Sensitivity (mV/g)	Acceleration (a)	Frequency Hz
2 cm vertical drop	vertical				
2 cm vertical drop	horizontal				
floor vibration	vertical				
floor vibration	horizontal				
mechanical vibration	any angle				
mechanical vibration	rotate 90°				

SELF CHECK

1. Sketch and label the five major components of a single-ended compression accelerometer, including the critical axis, and describe in a sentence or two how it works.

2. Sketch and label the three major components of a shear-design accelerometer, its critical axis, and give a one- or two-sentence description of how it works.

3. Sketch the voltage-equivalent circuit of an accelerometer, and label the signal source, capacitance, and resistance.

4. Name a type of accelerometer measuring device, and describe in one or two sentences its key characteristics.

5. An accelerometer with a sensitivity of 15 mV/g produces a voltage of 0.55 volts when in contact with an engine. What is the acceleration?

6. An accelerometer with a sensitivity of 12 mV/g is used to measure an acceleration of 2.6 g. What voltage is produced?

7. An oscilloscope is adjusted so that 6 divisions on the oscilloscope corresponds to a voltage of 50 mV. What is the voltage of a signal that has a pulse height of 5 divisions?

8. What is the sensitivity of an accelerometer that produces a voltage of 10 mV when measuring an acceleration of 0.15 g?

9. If an accelerometer has the same sensitivity as that of the one in problem 8 and is used to measure an accelerometer of 25 m/s^2, what voltage is produced?

10F1

BOURDON TUBES

The bourdon tube is a mechanical transducer used to measure pressure. The measurement of pressure is one of the most important functions in the field of instrumentation and control, not only because of the pressure measurement itself, but also because such measurement indicates other important parameters, such as temperature, liquid level, and flow. The bourdon tube pressure element transforms fluid energy or motion into detectable pressure changes on a readout device.

This module discusses the design and operation of the bourdon tube and, briefly, of the bellows and the diaphragm pressure transducers. In the laboratory, the student will operate and calibrate a bourdon tube.

OBJECTIVES

Upon completion of this module, the student should be able to:

- Compare the operation and application of five types of pressure elements used as transducers.
- Given a specific application, identify the appropriate pressure element.
- Given the appropriate equipment, operate and calibrate a bourdon tube.

DISCUSSION

The bourdon tube is a general type of pressure element, consisting of the simple "C" tube, helical, and spiral elements (Figure 1). Bellows and diaphragms are not classified as bourdon tubes, but are widely used as pressure transducers. Regardless of the type of transducer used, the basic principle of operation is the same: pressure applied to the tube causes a movement interpreted as a pressure measurement. A bourdon tube obeys Hooke's law, which states that within elastic limits, the free end of a bourdon tube will move in proportion to the fluid pressure applied. Thus, fluid energy is transformed to the potential energy of a spring-like element used to determine pressure.

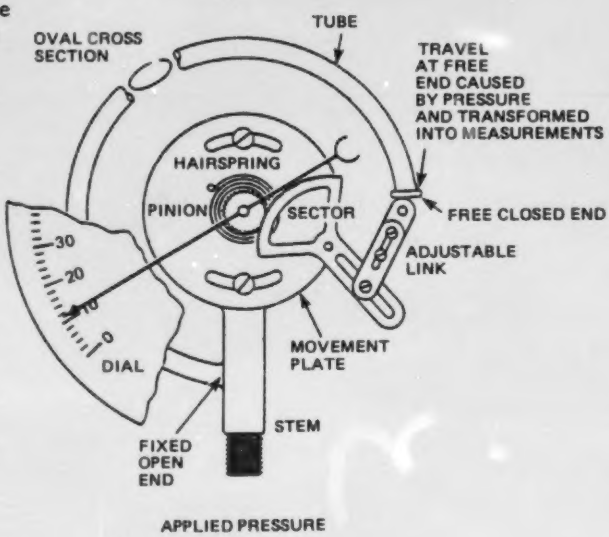

TUBE
OVAL CROSS SECTION
TRAVEL AT FREE END CAUSED BY PRESSURE AND TRANSFORMED INTO MEASUREMENTS
HAIRSPRING
PINION
SECTOR
FREE CLOSED END
ADJUSTABLE LINK
30
20
10
0
DIAL
MOVEMENT PLATE
STEM
FIXED OPEN END
APPLIED PRESSURE

FIG. 1

The bourdon C-tube shown in Figure 1 is the most common type of pressure element, the easiest to construct, and therefore, cheapest.

Pressure applied to the tube causes distortion in the shape of the device. Such distortion results in a small movement of the free end. This movement is rather small; so it is usually transformed and amplified by use of levers and gears to cause a rotation of a shaft. When a pointer is connected to the shaft, the shaft rotation causes the pointer to move along a graduated, calibrated scale. The relationship between pressure applied to the tube and movement of the tube depends upon tube construction and shape.

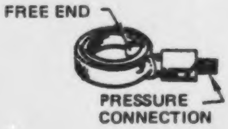

FREE END
PRESSURE CONNECTION

FIG. 2

Spiral (Figure 2) and helical (Figure 3) bourdon tubes operate on the same principle described above, but have the advantage of producing more travel of tip for a given applied pressure. The operation of a spiral element is the same as that of a "party favor," which tends to uncoil when pressure is applied to the device (Figure 4). The amount of uncoiling is proportional to the lung pressure applied; however, the spiral element does not uncoil completely like the party favor mentioned, because to do so would cause a distortion of the metal, and the element would not return to its original shape. One end of the spiral is mounted rigidly to a frame; the movement at the free end is transmitted to a pointer, which, in turn, moves along a calibrated scale to indicate pressure. Table 1 gives the pressure range that can be measured with different spiral-element materials.

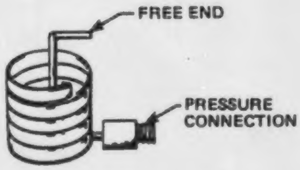

FREE END
PRESSURE CONNECTION

FIG. 3

PARTY FAVOR

FIG. 4

The helical element (Figure 3), instead of having coils in the same plane, is manufactured so that the coils are the same diameter and extend a vertical distance of one to five inches. The physical appearance resembles a coil spring.

Helical elements normally are made of thicker material and, consequently, can be used for higher pressure applications. Like the spiral, helical elements allow more movement at the tip than the C-tube and sometimes are used to drive an indicating device directly without the use of movement amplification by a link-and-lever system. Table 2 gives the pressure range for different helical element materials.

TABLE 1: PRESSURE RANGE OF SPIRAL ELEMENTS

Bronze	0-10 to 0-200 psi
Beryllium-Copper	0-10 to 0-200 psi
Ni-Span C Alloy	0-10 to 0-200 psi
Type 316 SS	0-12 to 0-200 psi

TABLE 2: PRESSURE RANGE OF HELICAL ELEMENTS

Bronze	0-201 to 0-400 psi
Beryllium-Copper	0-201 to 0-6000 psi
Ni-Span C Alloy	0-201 to 0-6000 psi
Type 316 SS	0-201 to 0-6000 psi
Heavy Duty Helical Elements:	Pressure Range
Type 316 SS	0-75 to 0-30,000 psi
	0-40,000 to 0-80,000 psi

All pressure elements operate on the energy transformation principle. Pressure applied to the element causes a movement. Since pressure applied through a distance represents fluid energy (recall, $\Delta W = p\Delta V$), the fluid energy is transformed into the mechanical energy of the indicating system. In this case, the mechanical energy is contained as potential energy in the deformed spiral or helix; therefore, the distortion is a direct measure of the pressure. Depending upon the type of material and physical construction of the element, it will offer an opposition to movement. For a given pressure, the opposition to movement and the pressure on the element are in a state of equilibrium. When the pressure changes, the element will move until this movement causes the element to build up an opposition to this movement. The operation is similar to a force applied to a spring and the opposition to movement by the spring. The spring moves until its force is equal to the applied force.

The force exerted by a pressure element is proportional to the applied pressure and the area of the element. A large-area pressure element is required for the measurement of extremely small pressures. For small pressure application, the diaphragm element is used (Figure 5). This type of energy transformer will produce a considerable amount of movement when a small pressure is applied. Presssure gages and transmitters using a diaphragm as an element can measure pressures as small as 0.30 pounds per square inch. Table 3 gives the pressure range for different diaphragm elements.

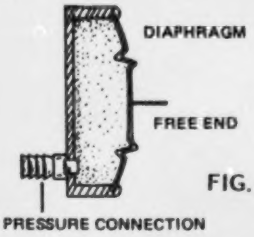


FIG. 5

TABLE 3: PRESSURE RANGE OF DIAPHRAGM ELEMENTS

Size	Pressure Ranges	
	Low	High
3-in, Cu-Ni-Mn	0-9 in H_2O (0-0.3psi)	0-40 in H_2O (0-1.4 psi)
2-in, Cu-Ni-Mn	0-20 in H_2O (0-0.7 psi)	0-5 psi (0-138.7 in H_2O)

The relationship between movement and pressure should be a direct proportionality for any type of pressure element. Ideally, equal increments of pressure should produce equal increments of movement at all points throughout the pressure range of the device. The bellows element is one of the most accurate pressure transducers because movement is more nearly proportional to pressure (Figures 6 and 7). A bellows element in operation resembles a small accordion. When pressure is applied to a bellows element, the side of the element expands; and when the pressure is removed, the side collapses to the original size. Table 4 gives the pressure range for different bellows-element materials.

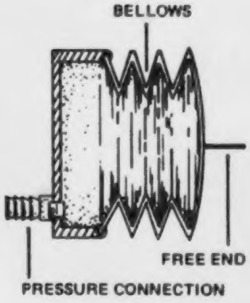


FIG. 6

TABLE 4: PRESSURE RANGE OF BELLOWS ELEMENTS

Type of Bellows Material	Pressure Ranges	
	Low	High
Brass	0-100 in H_2O (0-3.6 psi)	0-25 psi (0-693.3 in H_2O)
Stainless Steel (Type 316)	0-4.5 psi (0-124.8 in H 0)	0-29 psi (0-804.2 in H 0)

Pressure elements are used in pressure gages and transmitters, providing the driving force to operate pressure switches. Common applications of pressure switches include those used to start and stop water pumps and air compressors. Most pressure switches use a diaphragm or bellows, because a pressure applied to the large area of these elements can produce forces sufficient to operate the switch.

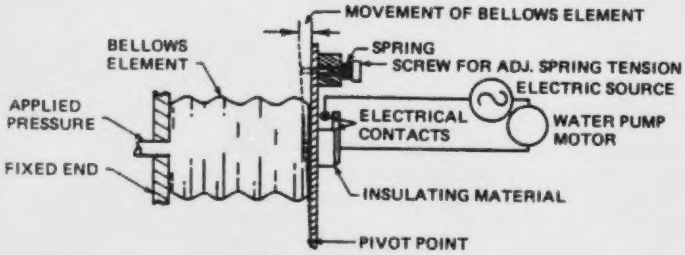

MOVEMENT OF BELLOWS ELEMENT
BELLOWS ELEMENT
SPRING
SCREW FOR ADJ. SPRING TENSION
ELECTRIC SOURCE
APPLIED PRESSURE
ELECTRICAL CONTACTS
WATER PUMP MOTOR
FIXED END
INSULATING MATERIAL
PIVOT POINT

FIG. 7

The switch (Figure 7) is in the de-energized position. When the pressure decreases, the bellows will contract, causing the electrical contact to close and start the motor to drive a pump. When the pump increases the pressure, the switch will again open, stopping the pump.

The adjustable spring is set to establish the best operating point.

SUMMARY

The bourdon tube is a transducer that transforms a pressure change into a needle-indicator deflection. The elastic element is a tube fixed at one end (open to accept applied pressure) but free at the other end (closed to allow displacement under the deforming action of the pressure difference across the tube walls). The most common design consists of a tube having an oval cross section and bent in a circular arc which straightens out slightly under pressure. By an almost frictionless linkage, the free end of the tube rotates a pointer over a calibrated scale to give a pressure indication.

OPERATION AND CALIBRATION OF A BOURDON TUBE

MATERIALS

0-10 lb (4 1/2" diameter) pressure gage using a C-tube as an energy transformer.
A 30-lb pressure source: a small, portable air compressor can be used. Caution: Do not use air pressure higher than 30 lb/square inch. Higher pressures can be dangerous. (Note: Many labs already are equipped with pressurized air ports.)
A "U" tube mercury manometer. (Caution: Care should be taken when working with mercury to avoid contact with skin and eyes. Mercury also is corrosive to gold jewelry.)
Variable pressure regulator.
Assorted pipe fittings and rubber hose.

All equipment needed for the exercises in this module is available from the Curtis Scientific Company, 4220 Jefferson Avenue, Houston, Texas 77001.

The manometer used in this module can be replaced with a previously-calibrated pressure gage. The standard practice, however, is to use a mercury manometer as a readable and accurate standard for calibration purposes.

PROCEDURE

The purpose of this procedure is to allow the student to observe the operation of a pressure element and to check the calibration of a pressure instrument, using a pressure element as an energy transformer.

1. Connect the equipment as shown in Figure 8.
2. Remove the glass, pointer, and dial from the pressure gage. This removal should expose the Bourdon C-tube. Observe the mechanical linkage by which the movement at the free end of the C-tube is converted to a rotary motion at the pointer hub.
3. In order to observe the operation of the C-tube during the test, do not replace the dial.

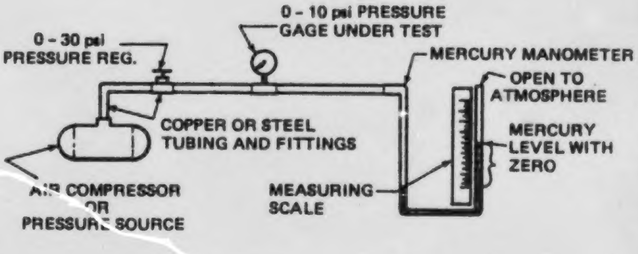

0 - 10 psi PRESSURE GAGE UNDER TEST
0 - 30 psi PRESSURE REG.
MERCURY MANOMETER
OPEN TO ATMOSPHERE
COPPER OR STEEL TUBING AND FITTINGS
MERCURY LEVEL WITH ZERO
AIR COMPRESSOR OR PRESSURE SOURCE
MEASURING SCALE

FIG. 8

With a soft lead pencil, mark the case of the gage in 10% increments corresponding to the dial markings.

4. With no pressure applied to the gage, replace the pointer on the zero pound mark.

5. Carefully adjust the pressure regulator until the manometer reads a pressure equal to 1 lb/square inch. (Note: To convert the reading on the mercury manometer to pounds per square inch, use the following conversion: 1 inch mercury = 0.4912 lb/in^2.) In Data Table, record the pressure reading on the pressure gage.

6. Repeat settings on the pressure regulator in increments of 10%. Record.

7. Calculate and record % error as indicated in Data Table.

DATA TABLE

% Reading on Gage	% Reading on Manometer	Error (%)*
0		
10		
20		
30		
40		
50		
60		
70		
80		
90		
100		

*Note: To calculate % error —

$$\%\ \text{error} = \frac{\text{manometer reading} - \text{gage reading}}{\text{manometer reading}} (100)$$

8. Make a plot of the readings from the Data Table, and determine if the gage is accurately measuring the pressure. The graph should be similar to the one shown in Figure 9, in which line A represents good calibrations. When all points

fall on a straight line 45° to both axes, the gage is calibrated properly.

Lines B represent zero point error — either high or low. The line is parallel to line A, but is offset by some constant amount.

Lines C represent an error caused by too much or too little span adjustment. The error increases as the reading progresses from the zero point.

If the calibration curve is other than line A, the pressure gage is not properly calibrated. In this case, proper calibration adjustment should be made. The adjustments are necessary to cause the movement of the pointer to correspond correctly to the movement of the C-tube or pressure transducer (Figure 9). The relationship of pressure and movement of the pressure element cannot be changed. This relationship is established by physical characteristics, such as the thickness and material type of element, when it is constructed. Movement of the pointer must correspond to the tip travel of the pressure element.

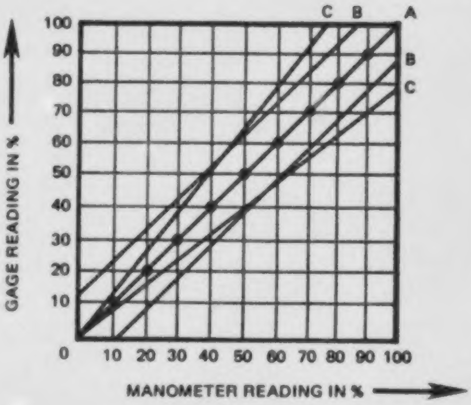


FIG. 9

A zero adjustment can be made by removing the pointer from the shaft attached to the pinion gear and by positioning the pointer to the 0 location on the scale. By loosening screw "s" and sliding the lever, the mechanical advantage of the device can be changed. This is a span adjustment.

9. Calibration procedure. (See Figure 10.)

   a. With zero pressure applied to the pressure gage, position the pointer to zero on the scale.

   b. Apply 100% pressure — 10 psi for a 10 psi gage — and note the gage reading. If too high, move the span adjustment in such a manner as to reduce the mechanical advantage of the device. If too low, increase the mechanical advantage.

   c. Manipulate the pressure regulator until zero pressure is applied to the gage. Position the pointer until it corresponds to the zero on the scale.

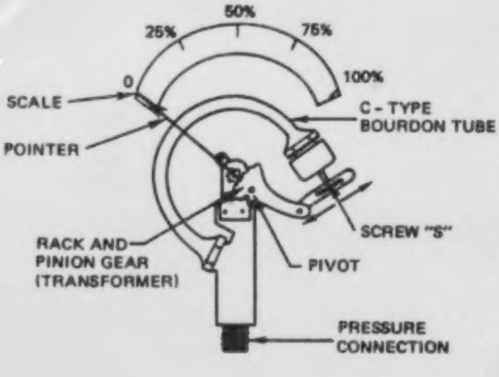


FIG. 10

d. Repeat Steps b and c until no further adjustment is necessary.

If the gage is new, the calibration should be good. To gain practice in calibration and to observe operation of the C-tube, uncalibrate the gage by making a random adjustment to zero and span, and go through the calibration procedures.

SELF CHECK

1. Discuss the operation and application of the following types of pressure elements used as transducers:

    Bourdon C-tube.
    Bourdon spiral tube.
    Bourdon helical tube.
    Diaphragm element.
    Bellows element.

2. Compare the operation and application of the Bourdon helical tube to that of the bellows element.

3. A pressure system operates at 4500 psi with a gas that corrodes type 316 SS tube. What types of helical element material are available to the system designed?

4. If a bellows moves 0.5 inches when the applied pressure is 10 psi, how much pressure is required to move the bellows 1 inch?

5. What is the total force applied to a diaphragm if the area of the diaphragm is 0.1 in^2 and the fluid pressure is 0.3 psi?

6. How much fluid energy is required to change the volume of an element by 0.001 in^3 if the pressure causing the volume change is 10 psi?

7. A force of 5N is applied to the fixed end of a bellows that has a cross-sectional area of 0.5 m^2. What pressure is exerted on the fixed end of this bellows?

8. Convert 10 psi to units of N/m^2.

10F3

DIFFERENTIAL PRESSURE TRANSDUCERS

Measurement of pressure is a common transducer application for both industrial and residential uses. Of practical importance is the distinction between absolute and gage pressure measurements, which involve different reference pressure levels. Differential-pressure transducers are typically employed in measurement of liquid levels and fluid flow rates.

OBJECTIVES

To complete this module, the student should be able to:

- Convert gage pressure to absolute pressure.
- Describe two types of differential-pressure transducers.
- Adjust and calibrate a differential-pressure transducer.
- Given a differential-pressure transducer, measure liquid level and fluid flow.

DISCUSSION

When barometric pressure is given as 760 mm of mercury or 14.7 psi, it is understood that this measurement is an absolute pressure, that is, referenced to zero mm of mercury or zero psi. This reference level would define a perfect vacuum. When one inflates a tire to 30 psi, on the other hand, the reference level is atmospheric pressure, or 14.7 psi. If atmospheric pressure should later change by 0.7 psi, this would represent a change in reference level. For example, if the reference level decreases by 0.7 psi and the same tire pressure is again measured with a tire gage, the gage would read 30.7 psi (30 + 0.7). The pressure of 30.7 psi is 0.7 psi higher than the initial pressure because the reference level was lowered by 0.7 psi. Absolute pressure inside the tire (compared to zero pressure) does not change, but measured pressure is different because the reference level changes.

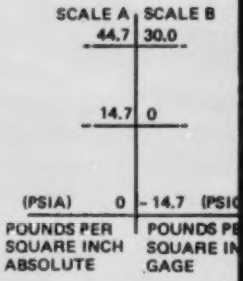


FIG. 1

Figure 1 illustrates two scales of pressure measurement. The measurement on scale A is referenced to absolute zero pressure (psia) or a vacuum. Scale B measurements are gage pressures (psig) and are referenced to atmospheric pressure, which is 14.7 pounds per square inch above zero absolute. It can be seen from the scale in Figure 1 that psia = psig + 14.7.

Measurements on both scales are in effect differential measurements because, as mentioned earlier, psia is the difference between some measured pressure and absolute zero pressure, while psig represents a value that is measured as a difference between some pressure and atmospheric pressure. Elements that are used as transducers to measure pressure produce a movement proportional to difference in pressure between the inside and outside of the device.

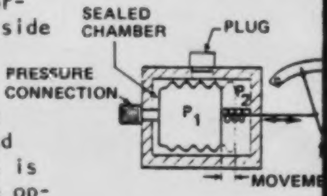


FIG. 2

When pressure is applied to a bellows (Figure 2), the pointer will move to the right. Movement will cease when a state of equilibrium exists between forces on the inside and outside of the bellows. Force on the inside of the bellows is caused by pressure applied to the bellows. What causes the opposing force on the outside of the bellows? A spring is commonly used, which helps establish pressure range of the bellows. Recall that force is the product of pressure and the area over which pressure is exerted. If pressure on the inside and outside changes by the same amount and in the same direction, there will be no resultant movement of the bellows and hence no change in indications.

Resultant force on the bellows can be shown by the following relationship:

$$F_1 = P_1A_1 \tag{1}$$

$$F_2 = P_2A_2 \tag{2}$$

where

F_1 = force on inside of bellows.
P_1 = pressure on inside of bellows.
A_1 = area of inside of bellows.
F_2 = force on outside of bellows.
P_2 = pressure on outside of bellows.
A_2 = area on outside of bellows.

By design of the bellows, $A_1 = A_2 = K$ and does not enter into the force relationship. Then:

$$F_1 = P_1K \tag{3}$$

and

$$F_2 = P_2K \tag{4}$$

By subtracting Equation 4 from Equation 3, we obtain:

$$F_1 - F_2 = (P_1 - P_2)K \tag{5}$$

From the above, it can be seen that difference in force that causes movement of the bellows is proportional to difference in pressure on each side of the pressure element.

For differential-pressure applications, diaphragms and bellows are most often used because of their ability to measure small differential pressure.

Construction of a differential-pressure transducer is such that pressure can be applied to both sides of the pressure element (Figure 3). One pressure will cause the indications to increase, while pressure applied to the other side of the element causes indication to decrease. The two ports through which the pressures are applied to the transducers are called the HIGH SIDE and LOW SIDE connections. Pressure applied to the HIGH SIDE connection causes the indication to increase, while pressure applied to the LOW SIDE causes a decrease in reading. Equal pressure applied to both connections causes no response from the instrument.

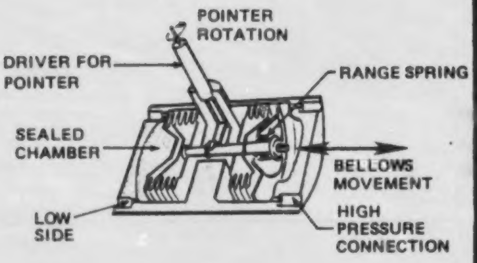

POINTER ROTATION
DRIVER FOR POINTER
RANGE SPRING
SEALED CHAMBER
BELLOWS MOVEMENT
LOW SIDE
HIGH PRESSURE CONNECTION

FIG. 3

TWO DIFFERENTIAL-PRESSURE TRANSDUCER APPLICATIONS

LEVEL MEASUREMENT

Recall from Module 10F1 that liquid level can be inferred from a pressure measurement that is caused by a hydrostatic head. Remember that a hydrostatic head is the pressure that a column of liquid exerts on the bottom of its container. An application for liquid level measurement by hydrostatic head is shown in Figure 4.

Relationship between hydrostatic head and liquid height can be expressed by the formulas:

$$\text{Pressure} = 0.433 \text{ psi/ft} \times (\text{specific gravity of fluid}) \times (\text{ft of head}).$$

Therefore,

$$\text{Level}_{(\text{ft})} = \frac{\text{pressure}}{0.433 \times \text{specific gravity}}.$$

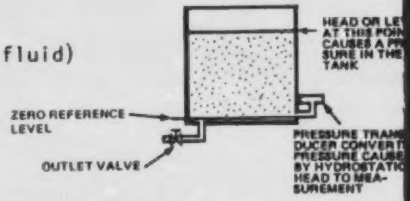

HEAD OR LE
AT THIS POIN
CAUSES A PR
SURE IN THE
TANK
ZERO REFERENCE LEVEL
OUTLET VALVE
PRESSURE TRAN
DUCER CONVERT
PRESSURE CAUSE
BY HYDROSTATIC
HEAD TO MEA-
SUREMENT

FIG. 4

This is an effective means of measuring level. Consider, however, the error caused by a change in pressure at the top of the tank. As with measurement of any variable, a response of the measuring device should be caused by a change in the measured variable. If the application of Figure 4 is to provide an accurate means of level measurement, it is necessary for pressure on the top of the liquid to remain at the same pressure as that on the outside of the pressure element used in the transducer, which would be atmospheric pressure. As a review of this application, refer to Figure 5, which shows that a change in

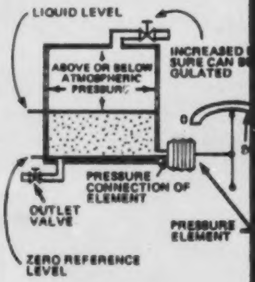

LIQUID LEVEL
ABOVE OR BELOW ATMOSPHERIC PRESSURE
INCREASED
SURE CAN B
GULATED
0
PRESSURE CONNECTION OF ELEMENT
OUTLET VALVE
PRESSURE ELEMENT
ZERO REFERENCE LEVEL

FIG. 5

pressure above the liquid level would cause serious error in measuring the liquid level.

When tank pressure is other than that on the outside of the bellows, the transducer responds and indicates a false change in level.

By using a differential-pressure transducer to measure level, the error caused by variations in static pressure inside the tank is eliminated. Pressure at the top of the tank is applied to one side of a differential-pressure instrument as in Figure 6, while the other side of the instrument is connected to the tank. The point at which the instrument is attached to the tank is the reference or zero level position. By observing the application shown in Figure 6, it can be seen that a fluctuation in pressure in the tank will change the force on the outside of the differential-pressure transducer. Note, however, that, while this is happening, pressure on the liquid causes a force of the same amount to be exerted on the inside of the bellows. Thus, the uncontrolled fluctuations in pressure in the tank cause equal forces to be applied to both the inside and outside of the bellows, resulting in no movement of the instrument. The hydrostatic head of the liquid is the only variable that will cause a movement of the measuring device.

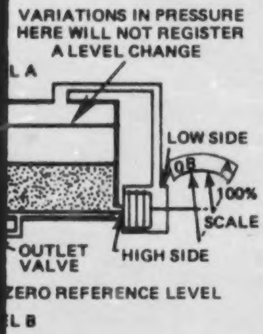

VARIATIONS IN PRESSURE HERE WILL NOT REGISTER A LEVEL CHANGE
L A
LOW SIDE
0
B
A
100%
SCALE
OUTLET VALVE
HIGH SIDE
ZERO REFERENCE LEVEL
L B

FIG. 6

FLUID FLOW MEASUREMENT

By referring to Module 10F2, it can be seen that fluid flow is proportional to the square root of the difference in pressure across a differential producer such as an orifice plate. Difference in pressure is measured by a differential-pressure transducer.

Figure 7 shows an application for measurement of flow with a differential-pressure transducer.

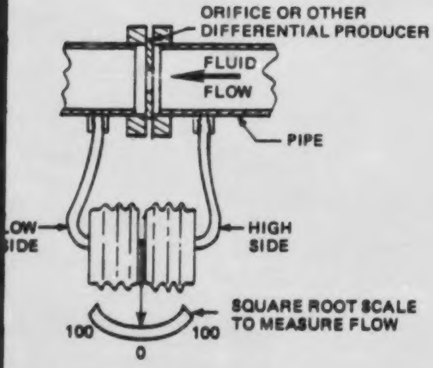

ORIFICE OR OTHER DIFFERENTIAL PRODUCER
FLUID FLOW
PIPE
OW SIDE
HIGH SIDE
SQUARE ROOT SCALE TO MEASURE FLOW
100
100
0

FIG. 7

LABORATORY

MATERIALS

Differential-pressure transducer with adjustable range.
Variable pressure regulator.
Air supply - 20 psig regulated.
Manometer with inches of water scale.

PROCEDURE

The purpose of this experiment is to enable the student to observe operation of a differential-pressure (d/P) transducer and to use standard test procedures to check calibration of the instrument.

1. Connect d/P instrument as shown in Figure 8.
2. Adjust air pressure regulator until manometer indicates zero. At this condition, atmospheric pressure is applied to both sides of d/P transducer.
3. Turn zero adjust screw on d/P transducer for an indication of zero differential pressure on scale.
4. Select a range in adjustable limits of instrument and adjust air regulator until manometer indicates selected full scale value.
5. If instrument reading is above or below full scale value, decrease or increase span as necessary to provide proper span on instrument. When performing step 5 for span adjustment, it will be necessary to make span adjustment and then reduce differential pressure to zero to observe result of span adjustment. This is true because a span adjustment provides multiplication of instrument's response.

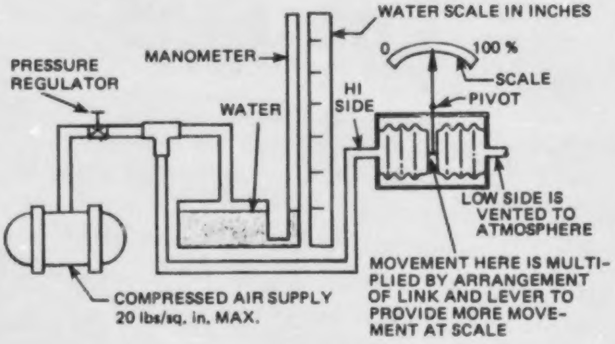

WATER SCALE IN INCHES
0
100 %
SCALE
PIVOT
MANOMETER
PRESSURE REGULATOR
WATER
HI SIDE
LOW SIDE IS VENTED TO ATMOSPHERE
MOVEMENT HERE IS MULTIPLIED BY ARRANGEMENT OF LINK AND LEVER TO PROVIDE MORE MOVEMENT AT SCALE
COMPRESSED AIR SUPPLY 20 lbs/sq. in. MAX.

FIG. 8

6. Return indicator on d/P transducer to zero by adjusting regulator until manometer indicates zero differential pressure.
7. Repeat step 3.
8. Repeat steps 3 through 7 until no further adjustment is necessary.
9. When zero and span adjustments have been made for proper indication of instrument at zero and one hundred percent of

d/P range, vary d/P in increments of 10% and list data in data table. To calculate % error, divide difference in indication between instrument and manometer by manometer indication and multiply by 100, or

$$(5) = \frac{[(2) - (3)]}{(3)} \times 100.$$

10. Draw a calibration curve by plotting manometer indication against instrument indications in Figure 9.

Note: Consult manufacturer's manual on specific instrument used for proper methods of zero and span adjustments as well as for calibration procedure.

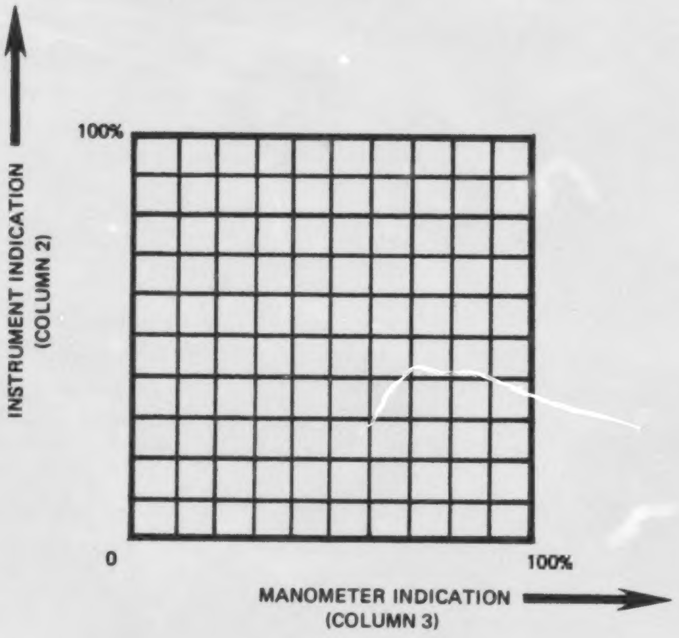


FIG. 9

DATA TABLE				
(1) Percent of Full Scale d/P	(2) Instrument Indication	(3) Manometer Indication	(4) Difference Between d/P Instrument and Manometer	(5) % Error
0				
10				
20				
30				
40				
50				
60				
70				
80				
90				
100				

SELF CHECK

1. Convert the following gage pressures to absolute pressures:

   3.0 psig to _________ psia.

   16.0 psig to _________ psia.

   152.0 psig to _________ psia.

   256.0 psig to _________ psia.

2. Describe two types of differential-pressure transducers.
3. Describe, in simple steps, calibration of a differential-pressure transducer.

4. Describe measurement of liquid level and fluid flow using a differential-pressure transducer.

5. A force of 10 lb is applied to the outside of a bellows and a force of 6 lb is applied to the inside of the bellows. If the resulting differential pressure is 20 lb/in^2, what is the area of the bellows?

6. How much pressure is exerted on the bottom of its container by a 3 foot high column of water?

7. A force of 20 lb is applied to the outside of a bellows and there is a force of 12 lb on the inside of the bellows. What is the pressure differential if the area of both the inside and the outside of the bellows is 6 in^2?

8. What head is required to produce a hydrostatic pressure of 6 psi at the bottom of a container of water? If atmospheric pressure is 14.7 lb/in^2 what is the total pressure at the bottom of the container?

10F4

FLUID-FLOW MEASURING DEVICES

Most systems applications depend upon the precise manipulation of fluid flow quantities to achieve a high quality of control. Before flow can be controlled accurately by automatic means, it must first be measured. For this reason, one must understand how to measure fluid flow. This module will cover flow measurement techniques with emphasis on those most commonly used.

OBJECTIVES

To complete this module, the student should be able to:

- Given the appropriate equipment, make measurements of the variables that relate to fluid flow velocity, calculate flow quantities, and discuss the relationships of the variables to the rate of flow.
- Assemble and operate flow measuring systems in the laboratory.

DISCUSSION

The ratelike quantity in the fluid systems is the amount of fluid flowing through a pipe, expressed as a volumetric or weight quantity per unit time. Such units are:

Gallons per minute
Barrels per day
Cubic feet per second
Pounds per hour

If the velocity of material moving through a pipe and the cross-sectional area of a given portion of that pipe are known, then the volume rate of flow can be determined from their product. If velocity is in ft/s and area in ft^2, then the volume rate is in units of $(ft/s)(ft^2) = ft^3/s$.

The standard approach to flow measurement is to use a portion of a pipe with known area and measure the time required for the fluid to travel through the specified length of pipe. The practice of flow measurement is, thus, a matter of measuring fluid velocity.

The usual approach to measuring fluid velocity incorporates a device which converts the kinetic energy of the moving material to a measurable signal. This is done with the help of devices which convert one form of energy to another.

Not only is it necessary to convert the kinetic energy to a measurable signal, but the signal must have an exact relationship to the flow rate. This relationship can be established by both analytical and empirical means.

HEAD TYPE FLOW MEASUREMENT

To understand flow measurement by head-type flowmeters, consider a fluid flowing at a steady rate through a pipe with a constriction placed in it. If the fluid is incompressible, the same volume of fluid must pass through every section of the pipe in the same time interval (Figure 1).

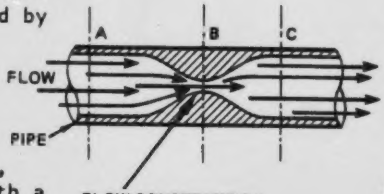


FIG. 1

Conforming to the laws of conservation of mass and energy, the mass entering any portion of pipe must equal mass leaving that portion during same time interval, and same is true for kinetic energy, as long as no energy is dissipated through friction. The fluid flow velocity is, therefore, increased in smaller section of pipe, resulting in an increase in kinetic energy per unit mass. In Concept Module 8-0, the principle of conservation of energy is applied to fluid flow, resulting in Bernoulli's equation. This fundamental equation relates in a general way the fluid velocity and pressure at different points in a pipe.

When a constriction is introduced in a pipe, a pressure difference appears across the constriction (called a "differential producer"). The pressure drop across the differential producer is usually measured in inches of water. The points in the pipe where the measurements are made are called taps and the location of the taps with respect to the differential producer is significant.

A basic application of flow measurement by differential pressure is shown in Figure 2, where a flat concentric orifice is flanged into the pipe to produce the differential pressure. The tap locations are fixed in the orifice flange at points one inch upstream and one inch downsteam from the orifice. The differential pressure is measured with a mercury manometer attached to the orifice flange (Figure 2). The pressure difference is proportional to the height difference in the two arms. At zero flow the mercury is at the same level on both sides of the manometer. The difference in levels h is greatest for the maximum flow.

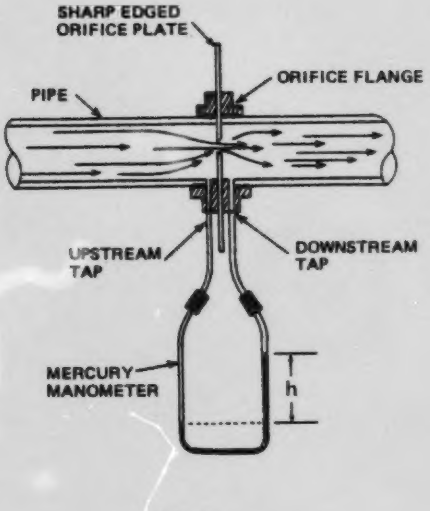


FIG. 2

The relationship of flow rate to differential height h, can be established empirically or described by theoretical formulas. The fact that such formulas exist is a major advantage of this simple type of flow measurement. Other more complicated flow measurement techniques must rely on empirical calibrations of the primary flow device.

A working equation has been derived to calculate the rate of flow through a differential pressure producer of known area and a pipe of known size at a given differential pressure. The working equation is:

$$Q = C_n\sqrt{\Delta p} \tag{1}$$

where

C_n = constant for particular flow situation.
Δp = differential pressure developed across producer.
Q = calculated flow rate.

For given flow conditions, values of C_n can be found in flow engineering books.

A magnetic flowmeter is another widely used method of flow measurement (Figure 3). The operation of a magnetic flowmeter is based on Faraday's law of induction, which states that the voltage induced in a conductor moving at right angles through a magnetic field is proportional to the velocity of the conductor moving through the field.

The moving conductor is the liquid and a magnetic filed is established at right angles to the pipe by magnetic coils. The voltage induced in the moving fluid is proportional to the fluid velocity and is detected by electrodes in contact with the fluid. This voltage is transmitted to a receiving instrument called a signal converter and is transformed to a scaled signal that represents a definite flow rate. For a particular meter, the strength of magnetic field is fixed as is the distance between the electrodes. The induced voltage is then purely a function of fluid velocity. Calibration of the magnetic flowmeter is established empirically by the manufacturer. The signal converter can be checked for accuracy by using a calibrator to simulate an electrical signal corresponding to a flow quantity.

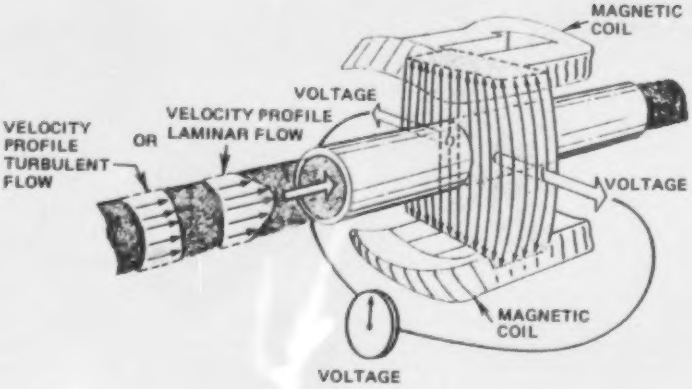

MAGNETIC COIL
VOLTAGE
VELOCITY PROFILE LAMINAR FLOW
OR
VELOCITY PROFILE TURBULENT FLOW
VOLTAGE
MAGNETIC COIL
VOLTAGE

FIG. 3

FLOW MEASUREMENT WITH A PLATE TYPE ORIFICE METER

MATERIALS

One orifice flange mounted in 2" pipe with 0.5" or 0.75" diameter orifice. This item can be manufactured in any machine shop capable of producing material to within 5/1000 tolerance or purchased from one of several reputable flow measuring companies.

Pump that can deliver a flow of about 10 gal/min.

Mercury "U" tube manometer of a differential pressure measuring device.

PROCEDURE

The purpose of this procedure is to establish a calibration curve for a plate type orifice meter.

1. Connect the orifice and manometer as shown in Figure 2 and establish a convenient flow rate to yield a measurable height differential, h.
2. Measure the flow by weight or volume per unit time in a convenient tank.
3. Record the value as a maximum flow rate in Data Table 1.
4. Reduce the flow to give a 10% reduction of height differential, h. Measure and record in Data Table 1 the flow rate at the lower height differential, h.
5. Repeat step 4 for eight different measurements each successively reduced in height differential by approximately 10%.
6. Plot each flow rate versus the square root of the corresponding height differential h in accordance with Equation 1. That is, plot Q versus $\sqrt{h}$ (Figure 4). Note: Since the pressure differential Δp is proportional to the height differential h, Equation 1 can equally well be stated as $Q = C_n' \sqrt{h}$, where C_n' is merely a constant different from C_n. The plot of Q versus $\sqrt{h}$

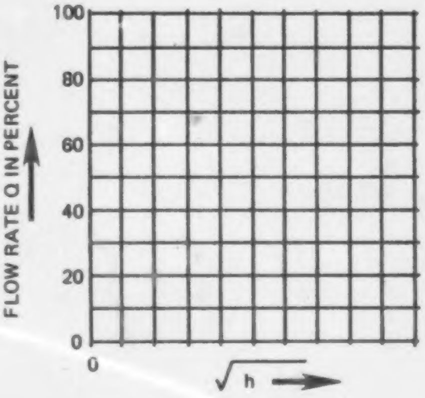


FIG. 4

should be linear and should be plotted on a graph (Figure 4).

7. From the plot in Figure 4, other flow rates for appropriate differential heights, h can be determined directly.

DATA TABLE 1	
Flow Rate in Gal/Min	Differential (h) in inches of water
1. (max)	(max)
2.	
3.	
4.	
5.	
6.	
7.	
8.	
9.	
10.	

FLOW MEASUREMENT WITH A TURBINE METER

MATERIALS

Flow nozzle of jet.
Impeller or fan blade mechanically connected to a small generator.
Voltmeter to measure output of generator.
Pump to produce a flow of about 10 gal/min.

PROCEDURE

The purpose of this procedure is to demonstrate the principle of flow measurement with a turbine meter.

When an impeller is placed in a pipe and is free to rotate, kinetic energy of the flowing fluid is converted to mechanical energy, causing the impeller to rotate. The speed of rotation is directly proportional to the fluid velocity. The rotary motion generates an electrical signal much like the moving fluid does in a magnetic flowmeter. This provides a very accurate means of flow measurement. It can be used for nonconducting fluids that cannot be measured by magnetic flowmeters. Calibration is done empirically by the manufacturer.

1. Arrange the water jet or flow nozzle in the end of a hose connected to the pump discharge.
2. Position the nozzle so a water stream emerging from it will start the impeller causing it to run the generator.
3. Measure the water flow and record a corresponding voltage output from the generator in Data Table 2.
4. Plot the generator output versus flow rate on a plot similar to Figure 4.

   This should produce a linear response of flow with respect to signal unless frictional losses in the impeller are significant. Thus, the relationship of signal to flow has been established.

DATA TABLE 2

Flow (Gal/Min)	Voltage
1.	
2.	
3.	
4.	
5.	
6.	
7.	
8.	
9.	
10.	

SELF CHECK

1. Given the appropriate equipment, make measurements of the variables that relate to fluid flow velocity, calculate flow quantities, and discuss the relationships of the variables to the rate of flow.

2. Assemble and operate flow measuring systems in the laboratory.

3. Express 20 gallons/minute in units of cubic feet per second.

4. The diameter of a portion of a pipe is 3 inches. If the fluid velocity through the pipe is 3 ft/s, what is the volume flow rate through the pipe? If the fluid is water (density of water = 1.94 slugs/ft^3) what is the mass flow rate through the pipe?

5. Water passes from a pipe of cross-sectional area 0.1 ft^2 to one of cross-sectional area 0.05 ft^2. If the flow velocity through the first pipe is 2 ft/s what is the flow velocity through the second pipe? (Hint: use the fact that the volume flow rate through the two pipes must be the same.)

6. The differential pressure across a constriction in a pipe is 150 lb/ft^2. What is C_n for this constriction if the volume flow rate through the pipe is 6 ft^3/s?

7. C_n for a particular pipe constriction is 0.2. (Q in ft^3/s and Δp in lb/ft^2) What is the volume flow rate through the pipe when the differential pressure across the constriction is 50 lb/ft^2?

10F5

VACUUM GAGES

Vacuum technology, used in science and industry, deals with gases at pressures below one atmosphere. Modern techniques can produce ultrahigh vacuum, with pressures less than 10^{-9} millimeters of mercury (compared to the pressure of 760 millimeters of mercury at one atmosphere). Even in such a vacuum, there are many millions of gas molecules per cubic centimeter.

This module examines a variety of different gages used to measure vacuum. Some gages, such as mercury manometers, respond to the pressure exerted by the gas. Other gages, such as the thermocouple gage, depend upon the change of a physical property of the gas as pressure decreases. In the laboratory, the student will use different types of vacuum gages to measure gas pressure in a vacuum system as the gas pressure is reduced below atmospheric pressure.

OBJECTIVES

Upon completion of this module, the student should be able to:

- State the pressure ranges implied by the following terms:

  Rough vacuum.
  Medium vacuum.
  High vacuum.
  Very high vacuum.
  Ultrahigh vacuum.

- Given the necessary information, calculate the mean free path of a gas molecule.

- Describe qualitatively how the following phenomena vary as gas pressure varies: thermal conductivity and impact ionization.

- Describe the operation of the following types of vacuum gages:

  Mechanical gage.
  Mercury manometer.
  McLeod gage.
  Thermocouple gage.
  Pirani (resistance) gage.
  Ionization gage.

- State the approximate range of pressure for which each of the above types of gages is useful.

- Given the appropriate equipment, measure the pressure in a vacuum system when the pressure is reduced below atmospheric pressure.

DISCUSSION

This module and Application Module 9F5 are two related modules concerned with vacuum technology. In this module, methods for measuring vacuum are emphasized; in Module 9F5, methods for producing it are emphasized.

The word "vacuum" means "empty space." But no vacuum is completely empty: There are always many gas molecules present within a so-called vacuum. As used here, the word "vacuum" refers to a space filled with a gas at a pressure less than one atmosphere.

According to this definition, a vacuum can be produced rather easily; for example, by a human being sucking on a straw, or by an octopus using the suction cups on its arms. Vacuum systems used in science and industry employ chambers (usually of metal) evacuated by pumps. Pressures that vary far below atmospheric pressure can be attained in such systems. This is the regime of vacuum technology addressed in this module.

A discussion of the units employed in the measurement of vacuum now is in order. The pressure exerted by a gas at one atmosphere pressure will support a column of mercury 760 millimeters high. As the pressure decreases, the height of the column that can be supported by the gas pressure decreases. Thus, gas pressures usually are expressed in the unit "millimeters of mercury," which refers to the height of the column of mercury that could be supported by the gas. Mercury manometers indicate pressure directly from the height of mercury in a tube. As pressures become very low (much less than 1 millimeter of mercury), direct observation of the height of a mercury column becomes impossible. Yet the unit "millimeter of mercury" still is used at very low pressures, even down to 10^{-10} mm of mercury and below.

Because the unit "millimeter of mercury" is cumbersome to use, it has been replaced in common use by the unit "torr," in honor of Torricelli, who invented the mercury barometer. Although there are some differences in the exact definition, for practical purposes one torr is equal to one millimeter of mercury. The torr unit will be used throughout the remainder of this module.

Vacuum technology covers a wide range of pressures, denoted by various terms. Table 1 summarizes some of the terminology commonly used. Slighly different definitions of the terms are used in different sources, but the tabulation here shows the general classification for different levels of vacuum. Of course, different measurement techniques are used in the various regimes. The lower limit to the range of ultrahigh vacuum continually decreases as vacuum technology progresses.

TABLE 1. VARIOUS LEVELS OF VACUUM

Level of Vacuum	Pressure Range (torr)
Rough vacuum	760 - 1
Medium vacuum	1 - 10^{-3}
High vacuum	10^{-3} - 10^{-6}
Very high vacuum	10^{-6} - 10^{-9}
Ultrahigh vacuum	below 10^{-9}

The pressure is related to the density of molecules in the gas by Equation 1.

$$p = nkT \tag{1}$$

where

p = pressure in units of force per unit area (in dyne/cm^2).
n = number of gas molecules per cm^3.
k = Boltzmann's constant, numerically equal to 1.38×10^{-16} erg/K.
T = absolute temperature K.

The relationship between pressure in dynes and torrs is expressed in Equation 2.

$$1 \text{ dyne/cm}^2 = 7.50 \times 10^{-4} \text{ torr}$$

or

$$1 \text{ torr} = 1333 \text{ dyne/cm}^2 \tag{2}$$

Equations 1 and 2 are employed in Example A to solve a practical problem.

EXAMPLE A: DENSITY OF GAS MOLECULES

Given: In a system, there is a gas pressure of 10^{-9} torr and a temperature of 300 K.

Find: The number of gas molecules per unit volume.

Solution: According to Equation 2, 10^{-9} torr is equal to 1.333×10^{-6} dyne/cm^2, which yields the following when substituted into Equation 1:

$$p = nkT$$

$$1.333 \times 10^{-6} = nkT$$

$$1.333 \times 10^{-6} = (n)(1.38 \times 10^{-16})(300)$$

$$n = (1.33 \times 10^{-6})/(1.38 \times 10^{-16})(300)$$

$$n = 3.22 \times 10^{7}/\text{cm}^{3}$$

Thus, even in an ultrahigh vacuum, there can be more than 30 million gas molecules in a cubic centimeter of space.

The composition of gas changes as pressure decreases. At pressures from 760 torr down through the medium vacuum range, gas is mainly residual air and has a chemical composition similar to that of air, about 80% nitrogen and 20% oxygen. As pressure drops to the high-vacuum range, the residual air is gone; and the gas in the system is that released by the surfaces in the system. In the high-vacuum range, the gas may be mostly water vapor, which arises from the release of water vapor absorbed on the surfaces. In the ultrahigh vacuum regime, the residual gas may be mostly hydrogen, coming from permeation of hydrogen entrapped within the materials from which the vacuum system is constructed. This change in composition is important for pressure measurement because different methods of measurement have varying sensitivities for different gases.

An important concept in vacuum technology is the "mean free path" of a gas molecule. The molecules in a gas are in random motion, colliding with each other and changing their direction of motion after each collision. The "mean free path" of a gas molecule is defined as "the average distance that a gas molecule travels between two successive collisions." Of course, as gas pressure decreases, the molecules are farther apart, and the mean free path increases. The mean free path is given by Equation 3.

$$\lambda = \frac{C_g}{n} \tag{3}$$

where

λ = mean free path.
C_g = constant, characteristic of the particular gas.
n = density of gas molecules.

According to Equation 1, the mean free path is inversely proportional to the pressure.

For air, C_g has the value of 1.66×10^{14} cm^{-2}. The value of C_g will be somewhat different for other gases, being higher for light monatomic gases, such as He, and slightly lower for polyatomic gases, such as CO_2. However, C_g does not vary greatly between different gases. Table 2 gives quantitative values of λ for air. At pressures near one atmosphere, the mean free path is very short. As pressure drops, the mean free path becomes larger and can be very large under conditions of very high vacuum or ultrahigh vacuum.

TABLE 2. MEAN FREE PATH (AIR AT 25°C)

Pressure (torr)	Mean Free Path λ (cm)
760	6.7×10^{-6}
1	5.1×10^{-3}
10^{-3}	5.1
10^{-6}	5.1×10^{3}
10^{-9}	5.1×10^{6}

As the mean free path becomes large as pressure decreases, it eventually becomes larger than the dimensions of the vacuum chamber. Then a gas molecule will make most of its collisions with the walls of the chamber and infrequently will collide with another gas molecule. It will travel through the vacuum chamber from wall-to-wall without encountering another gas molecule.

In order for vacuum technology to be used in practical applications, the level of the vacuum must be measured. There are many different methods used to measure gas pressure in a vacuum system. Because the pressures of interest in vacuum systems cover such a broad range, no single method is applicable for all pressure ranges. Different types of gages are useful in different ranges of pressure, and different gages may have special advantages or disadvantages for specific applications. Many different types of gages, therefore, are used in vacuum technology. The discussion below will describe the operation of some of the most frequently-used gages.

Some gages respond to the total pressure exerted by the gas. Such gages include the mechanical gage, the mercury manometer, and the McLeod gage. These gages generally are useful only in higher ranges of pressure and do not always offer continuous recording of pressure.

Other gages operate by sensing some gas property, such as thermal conductivity, that changes with pressure. Such gages record continuously and are useful in lower ranges of pressure. The gages of this type that will be discussed include the thermocouple gage, the Pirani (or resistance) gage, and the ionization gage.

THE MECHANICAL GAGE

In a mechanical diaphragm gage, a flexible metal diaphragm separates two chambers. One chamber is connected to the vacuum system whose pressure is to be measured; the other is evacuated to low pressure. As the pressure in the vacuum chamber changes, the diaphragm flexes slightly because of the varying air pressure on one side. The diaphragm is linked mechanically to a meter, so that the flexure of the diaphragm is displayed directly on the meter as a pressure. Often such gages are constructed as circular meters covering a specific range of pressure, such as 0-760 torr, or 0-20 torr.

Of course, at low pressures such gages lose sensitivity. The practical range covered by a gage in which the diaphragm simply is linked mechanically to the pointer in the meter is perhaps 0.1 torr to 760 torr. More elaborate means of coupling the motion of the diaphragm to the meter are possible; for example, optical or electrical coupling. With these methods, the range of operation of such a gage can be extended to lower pressures.

MERCURY MANOMETER

A manometer consists of a U-shaped glass tube filled with a liquid, usually mercury. One end is connected to the vacuum system, as indicated by Figure 1. The other end may be open to the air. As the pressure in the vacuum system is reduced, the level of mercury rises in the tube attached to the vacuum system. The difference in the height Δh between the levels in the two arms is proportional to the difference in pressure between the two arms. Thus, measurement of the height Δh (read from the graduated scale mounted next to the U-tube) gives a measure of the pressure in the vacuum system. The pressure p is given by Equation 4.

$$p = p_a - \Delta h \qquad (4)$$

where

p = pressure in vacuum system in torr (mm of Hg).
p_a = ambient air pressure in torr (mm of Hg).
Δh = differential height.

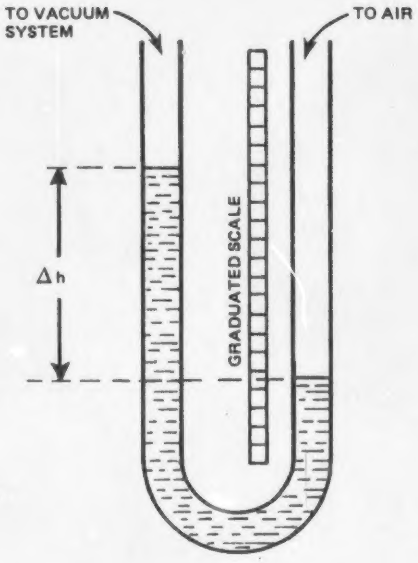


FIG. 1

The differential pressure between air and vacuum is sufficient to support a column of mercury 760 mm high, which, as described before, leads to the definition of the unit of pressure as a "millimeter of mercury," with ambient air having a pressure of 760 mm of Hg.

The range over which the mercury manometer is useful is determined by its ability to measure the height Δh, although by eye, one can measure it to a precision of about 0.1 mm. Thus, if the pressure were 1 torr, the error in the pressure would be about 10% because of the uncertainty in measurement of Δh. At lower pressures the percentage error would be larger. The mercury manometer, therefore, is useful for pressure measurements over the range of from 1 torr up to 760 torr. Its range could be extended to lower values by use of a microscope with which to meausre the level.

Over the pressure range for which it is applicable, the mercury manometer provides a simple, easy-to-use gage, reading directly without any need of calibration.

THE McLEOD GAGE

The McLeod gage also involves measurement of the height of a column of mercury in a glass tube. It employs compression of the gas by a known amount and extends the range of measurement below the range covered by the mercury manometer.

The basic structure of the McLeod gage is illustrated in Figure 2. It consists of a glass bulb with a capillary tube on the top. The bulb is connected to the vacuum system through a side arm that has a capillary section in parallel with the main tube. This capillary has the same diameter as the capillary on top of the bulb. There also is a means for raising and lowering the level of mercury in the system, shown here as an "auxiliary pump" connection.

When the mercury is lowered far enough that the bulb is connected to the vacuum system through the side arm, the pressure *p* in the bulb will be the same as the pressure in the vacuum system. Then the mercury level is raised so that the bulb is cut off from the side arm, and the gas is compressed into the capillary at the top of the bulb.

Since the two capillaries have the same bore size, surface-tension effects are canceled out. The difference in level of mercury in the two capillaries

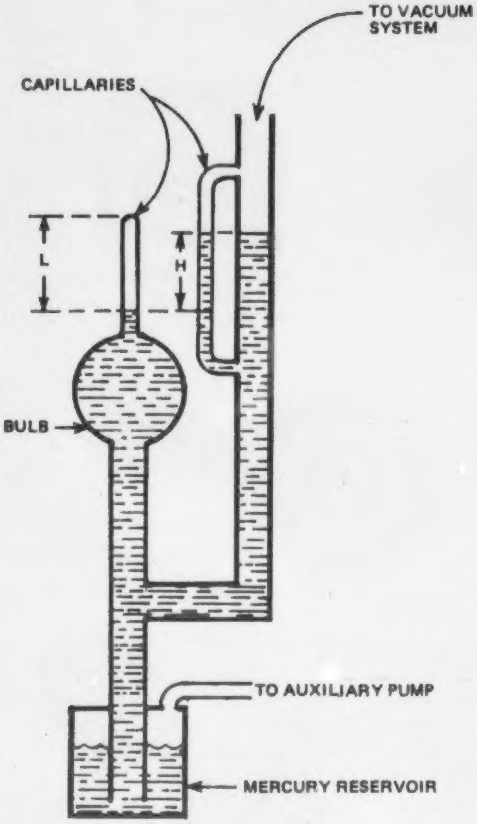


FIG. 2

arises from the pressure difference due to compression. The pressure p in the vacuum system is related to the geometrical factors of the gage by Equation 5.

$$p = \frac{Ah_1h_2}{(V - Ah_2)} \qquad (5)$$

where

p = pressure of vacuum system.
A = cross-sectional area of either capillary.
h_1 = height of mercury in capillary attached to side arm.
h_2 = height of mercury in capillary attached to bulb.
V = volume of bulb.

For a known gage construction, the values of A and V will be known, and the pressure measurement may be accomplished by measuring h_1 and h_2. If one allows the capillary to fill a constant, predetermined level, h_2 will be a known constant value; and one will have to read only h_1 from the graduated scale.

The use of compression allows a small pressure to be translated into a larger difference in height, as compared to the mercury manometer; therefore, the McLeod gage is useful at lower values of pressure, down to values as low as 10^{-6} torr. Because the gas in the bulb must be compressed, the gage cannot be used at pressures near atmospheric pressure. The upper limit of the useful range of pressure for the McLeod gage is approximately 100 torr.

The McLeod gage, therefore, covers a wide range of pressures and is a direct-reading instrument that responds directly to pressure. Thus, it often is used as a calibration device to calibrate other gages such as the thermocouples, Pirani, and ionization gages to be described later.

The McLeod gage has the disadvantage that it does not respond to condensable vapors such as water vapor. It also is somewhat cumbersome to use and does not provide continuous readings. Therefore, after calibration, other gages, such as the thermocouple gage, generally are employed for routine pressure monitoring.

THE THERMOCOUPLE GAGE

The operation of the thermocouple gage depends upon the cooling of a heated wire by thermal conduction through a gas. Thermal conductivity is measured by the extent to which a substance carries heat into its interior. The thermal conductivity k of a gas at temperature T is given by Equation 6.

$$k_t = \left(\frac{8\ m\ k\ T}{9\pi}\right) c\ n\ \lambda \qquad (6)$$

where

k_t = thermal conductivity of a gas at temperature T.
m = mass of gas molecules.
k = Botzmann's constant (as previously defined).
T = temperaure of as in K.
c = specific heat of the gas.
n = density of gas molecules.
λ = mean free path.

According to Equation 3, the mean free path is inversely proportional to n, so that the thermal conductivity will be independent of gas pressure. This relationship will be true for the range of gas pressures for which the mean free path is reasonably small. But as the pressure becomes low enough, the mean free path becomes comparable to the dimensions of the vacuum chamber. Then the mean free path cannot continue to increase with decreasing pressure, but remains constant, approximately equal to the size of the chamber. Then, as pressure decreases still further, thermal conductivity decreases, which, in turn, means that the heated wire cannot have heat removed from it by the gas as effectively as at higher pressures, and that the temperature of the wire increases. The temperature of the wire is measured by a thermocouple in contact with it.

From the description of thermal conductivity above, the thermocouple gage is useful over a limited range of pressures. At high pressures, thermal conductivity of the gas is independent of pressure, and the gage does not respond to changing pressure.

Thermocouple gages are available commercially as packaged devices designed to be inserted into the vacuum system, with electrical leads emerging from a feed-through. Two leads are used to monitor the voltage from the thermocouple and are used as the input to a calibrated meter, from which the pressure can be read directly. Two other leads are utilized to supply a calibrated electrical current that heats the wire. The current must be held at a constant value. The meter for a thermocouple gage usually is accompanied by a separate meter to monitor the current and to ensure that it remains constant.

The range of pressure covered by the thermocouple gage is limited by the physical phenomena described above. The thermocouple gage is useful over approximately the range from 10^{-4} torr to 10 torr. At higher pressures, thermal conductivity of the gas is independent of pressure, as earlier described. At lower pressures, thermal conductivity becomes so low that the gage is not sensitive to further changes in thermal conductivity.

THE PIRANI GAGE (OR RESISTANCE GAGE)

The Priani gage, or resistance gage, also depends upon the change of thermal conductivity of the gas with pressure. The gage has an electrically-heated wire cooled by the gas, and temperature of the wire depends upon gas pressure. The electrical resistance of the wire depends upon its temperature. Thus, measurement of the resistance of the wire will give a measurement of the gas pressure.

Often the measurement is performed with a Wheatstone bridge circuit as illustrated in Figure 3. The resistances R_2 and R_4 are fixed. The value of the variable resistor R_3 is changed until no current flows through the galvanometer G. Then the resistance R_1 of the gage is determined from Equation 7.

$$R_1 = \frac{R_2 R_3}{R_4} \tag{7}$$

In order to calibrate the response of the gage, one must calibrate the response of the resistance R_1 against some other gage; for example, a McLeod gage.

The Pirani gage is useful over the pressure range of about 10^{-4} torr to 10 torr, for the same reasons as were discussed for the thermocouple gage.

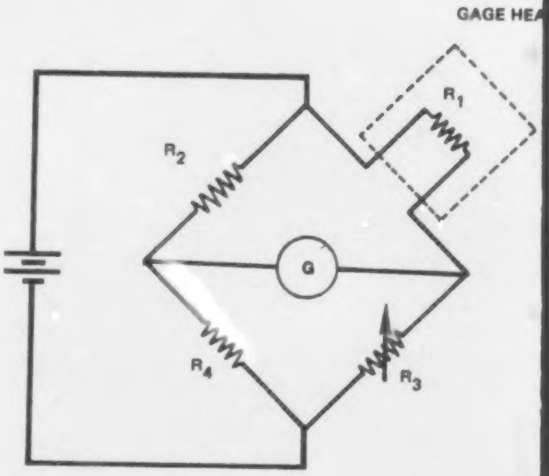


FIG. 3

THE IONIZATION GAGE

There is a variety of different types of ionization gages in use. We shall describe only the so-called "hot cathode ionization gage." The structure of this gage is illustrated in Figure 4. The fine-wire filament is heated electrically. Electrons are driven off the surface of heated metals in a phenomenon called "thermionic emission"; therefore, the heated filament serves as a source of free electrons.

The filament is surrounded by a wire grid held

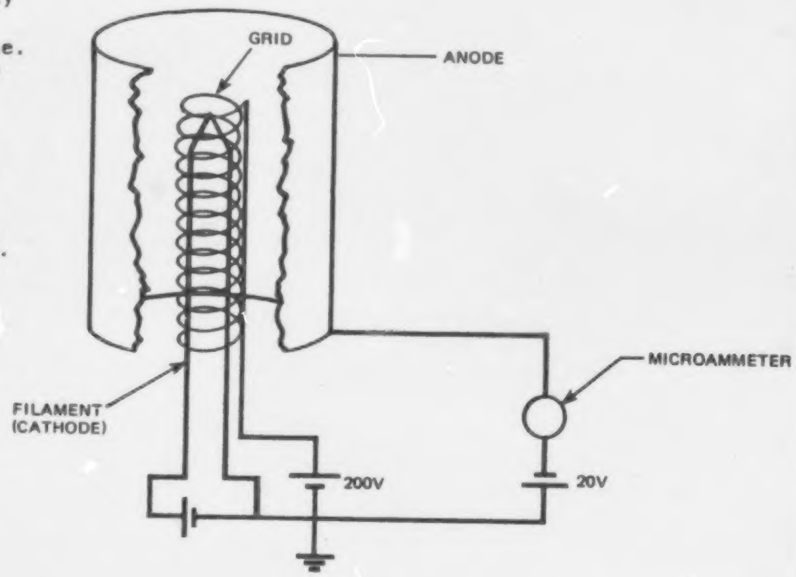


FIG. 4

at a potential of, perhaps, 200 V above that of the cathode. Thus, the electrons are accelerated toward the grid, depicted in Figure 4 as a wire helix surrounding the filament. The electrons acquire a large kinetic energy and may collide with gas molecules and ionize them, producing free positive ions. The free positive ions then travel to a collecting anode, shown in Figure 4 as a solid cylindrical shell surrounding the grid and filament. The anode is held at an electrical potential of about 20 V below that of the filament. Thus, electrons from the filament do not reach it, although the positive ions are drawn to it. This process results in an electrical ion current which can be read on the microammeter.

The rate of ion production is proportional to the density of gas molecules. The more gas molecules present, the more likely that an electron will collide with one and produce a free ion. The positive ion current, therefore, is proportional to the gas pressure, as expressed by Equation 8.

$$I_p = G\, I_e\, p \tag{8}$$

where

I_p = ion current.
G = gage sensitivity.
I_e = electron current from filament.
p = gas pressure.

The factor G depends upon the nature of the gas in the chamber and upon the applied voltages. To calculate G is very difficult. The value of G usually is determined by a calibration. Ionization gages must be calibrated against other gages, often McLeod gages, where their respective ranges overlap.

Hot cathode ionization gages are useful over the pressure range from about 10^{-8} torr to 10^{-3} torr. At higher pressures, the heated filament would burn out in the residual air; therefore, one must ensure that the ionization gage is not turned on until the pressure is reduced to a safe level.

The lower limit to the useful pressure range is determined by the production of X-rays by the electrons. The X-rays strike the anode and produce a spurious background current. There are additional variations of the ionization gage, which allow use of the ionization gage to measure pressures below 10^{-8} torr.

SUMMARY

This module has described the operation of several different types of gages capable of measuring the pressure of gas in a vacuum system. Table 3 summarizes the ranges of pressures over which these gages are useful. Although the gages described here include some of the more important ones, there are other types, some of which are capable of measuring pressures down to 10^{-14} torr, or even below. The values given in Table 3 must be considered as approximate ranges. Variations of some of the devices have been developed to operate at pressures outside these limits. The purpose of Table 3 is to give a general indication of the pressure range in which a given gage is most useful.

TABLE 3. USEFUL RANGES FOR VARIOUS PRESSURE GAGES

Gage	Useful Range (torr)
Mechanical Diaphragm	0.1 - 760
Mercury Manometer	1 - 760
McLeod	10^{-6} - 100
Thermocouple	10^{-4} - 10
Pirani	10^{-4} - 10
Hot Cathode Ionization	10^{-8} - 10^{-3}

MEASUREMENTS WITH VACUUM GAGES

MATERIALS

Vacuum System, including:

1. Mechanical pump (e.g., Welch Scientific).
2. Vacuum gages; at least two required and at least 3 desirable; for example, the following is a suitable combination:
   a. Circular mechanical gage (e.g., Wallace and Tierman Co.).
   b. McLeod gage (e.g., Bendix Co.).
   c. Thermocouple gage (e.g., Veeco).

3. Valve between pump and vacuum system.
4. Flanges and couplings, as needed.

PROCEDURE

The student will learn to use pressure gages by employing at least two different types of gages to measure the pressure in a vacuum system as the pressure is reduced below atmospheric pressure. The vacuum system should be set up beforehand since to require the student also to set up the system is beyond the scope of this module. The vacuum system should incorporate at least two different gages; three different types would be preferable. These combined gages should cover a range of pressures from one atmosphere (760 torr) down to a suitably low value. These procedures will assume that a mechanical gage, a McLeod gage, and a thermocouple gage are available. The procedures can be modified to employ a different set of gages.

1. Close valve and turn on vacuum pump.

2. Open valve, and observe decreasing pressure indicated on mechanical gage. Record pressure in Data Table. When pressure has reached a value within range of McLeod gage, close valve to stop pumping.

3. Measure pressure with McLeod gage; record in Data Table. With reference to Figure 2, lower the level of mercury in

gage to below position of side arm. Then raise level so that gas in bulb is compressed into capillary on top of bulb. Mercury should be high enough so that it partially fills capillary. Measure the heights h_1 and h_2, as defined in Figure 2.

4. Calculate pressure, using Equation 5. Use values of V and A supplied by manufacturer of gage. Record pressure.

5. Compare pressure obtained in Step 4 to pressure indicated by mechanical gage.

6. Open valve and continue pumping. Pump until pressure reaches a value within range of thermocouple gage. Record pressure. This value may be checked by periodically stopping pumping and using McLeod gage to measure pressure, using Steps 3 and 4. When the pressure is low enough, proceed to next step.

7. Turn on thermocouple gage power supply. Use current-adjusting set screw to set calibrated current at value specified by manufacturer.

8. Read and record pressure on pressure scale. Notice that pressure is indicated on a logarithmic scale. Record pressure. Compare the value to that obtained using McLeod gage.

9. Open valve, and continue pumping. Observe decreasing readings on thermocouple gage scale. Check current periodically to ensure that it remains at proper value. When pressure has stopped decreasing, record the limiting value of pressure obtained.

10. How does this experiment demonstrate the principles examined in the Discussion?

DATA TABLE

Mechanical Gage Pressure	McLeod Gage			Thermocouple Gage	
	Measured Pressure		Calculated Pressure	Pressure	Limiting Pressure
	h_1	h_2			

SELF CHECK

1. State the pressure ranges implied by the following terms:

   Rough vacuum.
   Medium vacuum.
   High vacuum.
   Very high vacuum.
   Ultrahigh vacuum.

2. What is the mean free path of a molecule in air at a pressure of 10 torr and at a pressure of 10^{-11} torr?

3. Describe qualitatively how these phenomena vary as gas pressure varies:
   Thermal conductivity.
   Impact ionization.

4. Describe the operation of the following types of vacuum gage:

   Mechanical gage.
   Mercury manometer.
   McLeod gage.
   Thermocouple gage.
   Pirani (resistance) gage.
   Ionization gage.

5. State the appropriate range of pressure for which each of the above gages is useful.

6. An enclosed volume of gas is at a temperature of 300K and a pressure of 5 torr. The pressure is reduced by lowering the temperature. What is pressure of the gas when the temperature is 78K?

7. One arm of a mercury manometer is connected to a vacuum system. The mercury level in this arm is 400 mm higher than the mercury level in the other arm which is open to the air. If the air pressure is 1 atmosphere, what is the pressure in the vacuum system?

8. An ionization gage is used to measure the pressure in a vacuum system. It has been determined that when the ion current is 10 µA and the electron current from the filament is 2 mA, the pressure in the vacuum system is 10^{-5} torr. What is the gage sensitivity?

9. What is the volume of the bulb of a McLeod pressure gage if the height of mercury in the capillary attached to the side arm is 20 cm, the mercury level in the capillary attached to the bulb is 10 cm, and the pressure in the vacuum system

attached to the gage is 1 Torr? Assume that the cross-sectional area of the capillaries is 0.001 m^2.

10E1

METER MOVEMENTS

The heart of many electrical measuring instruments is the meter movement. Meter movements are transducers that use both electrical and mechanical energy. In a meter movement, an electrical current flows through a coil of wire to produce a magnetic field, which tends to align itself with the magnetic field of a permanent magnet. A pointer attached to the coil indicates the value of the current flowing through the coil.

This module will discuss the operation and applications of meter movements. In the laboratory, the student will construct a voltmeter and an ammeter.

OBJECTIVES

Upon completion of this module, the student should be able to:

- List three major types of meter movements and an application of each type.
- Draw and label a sketch that illustrates the major components of a stationary-magnet/moving-coil meter movement; and explain its operation.
- Draw and label diagrams of the circuits used in the following types of meters:

  Voltmeter.
  Ammeter.
  Ohmmeter.
- Explain, with the use of a schematic diagram, how the internal resistance of a meter movement can be determined.

- Given the appropriate equipment, construct and operate a single-scale voltmeter and a single-scale ammeter.

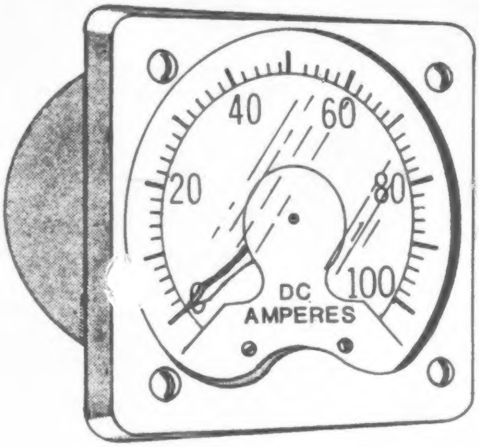
DISCUSSION

The ammeter and the voltmeter — two types of meter movement instruments used for switchboard and control applications — are depicted in Figure 1. They represent the many different electrical measuring instruments, the fundamental tools of the electrical industry.

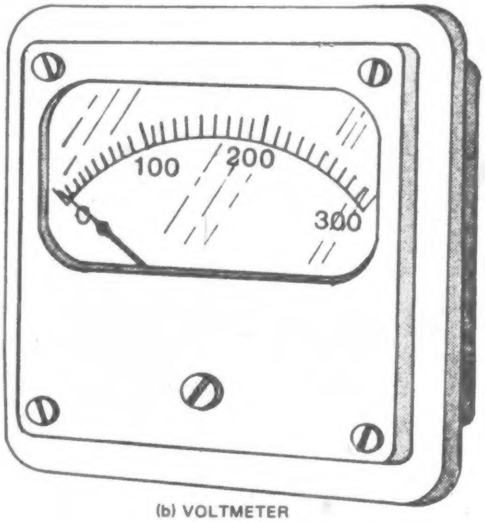
These instruments are essential to the evaluation of various electrical quantities under a variety of conditions. Thus, an understanding of the operating principles and the construction of instruments is an indispensible aid in their proper application.

All common electrical meter movements operate as the result of magnetic fields produced by current flow through a coil of wire. The following three major types are necessary components of instruments used for measurements under the many conditions encountered in modern technology:

- Stationary magnet. Moving-coil mechanisms are the most common. They consist of a stationary magnet and a coil free to rotate through an angle. These meters provide high sensitivity and rapid response because of the small mass that must be moved.
- Stationary coil. Moving-magnet mechanisms are less sensitive and slower to respond, but are more rugged. They sometimes are used in the measurement of large currents.
- Stationary coil. Moving-coil mechanisms are used in the measurement of a.c. quantities. The most common example of this type is the a.c. Wattmeter discussed in Module 6E1, "Electrical Power Measurement."



(a) AMMETER



(b) VOLTMETER

FIG. 1

All of these meter movements consist of a movable magnetic field suspended in a stationary magnetic field. The strength of one or both of the fields is proportional to current flow through the device. The remainder of this module is a description of the operation of the most common stationary-magnet/moving-coil mechanisms. The other types of meter movements operate according to the same principles.

STATIONARY-MAGNET / MOVING-COIL METERS

The stationary-magnet/moving-coil instrument illustrated in Figure 2 is used to measure undirectional currents, such as those supplied by a battery, rectifier, or d.c. generator. A strong permanent magnet is mounted rigidly in order that the magnetic field it provides will be of fixed position and polarity. The field is concentrated and made uniform across the air gap by soft iron pole pieces and a soft iron stationary core. Figure 3 is a more detailed diagram of the permanent magnet, pole pieces, and core.

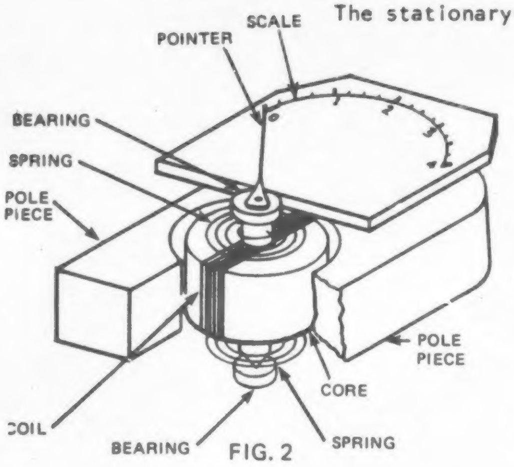


FIG. 2

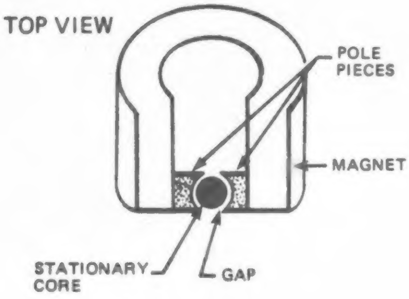


FIG. 3

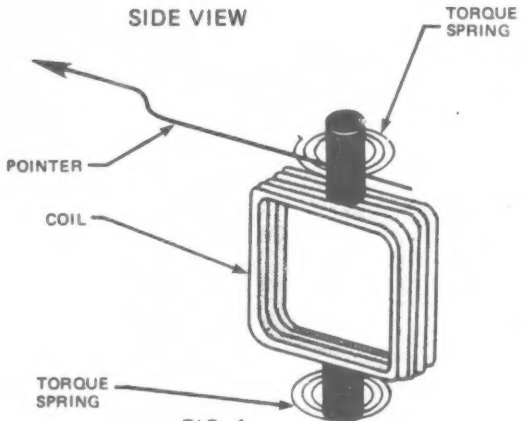


FIG. 4

A coil (Figure 4) pivoted to rotate on jeweled bearings is located in such a manner that it is free to move in the gap between the pole pieces and the core. The coil is connected to the electrical circuit through torque springs that provide the restoring force of the instrument and also carry current to the coil.

The moving element must be light in weight to respond quickly to changing current and be strong enough to resist the forces of overloads.

With no current flowing through the coil, no magnetic force acts upon the coil. It is held in the "zero" position by the coil springs. When a current flows, magnetic forces are produced on the movable coil (Figure 5). These forces result in a rotational torque that opposes the torque applied by the spring. The coil rotates to a new position, at which the two torques are of equal magnitude. The energy of the magnetic field is converted to potential energy of the coiled spring. A pointer appropriately attached to the coil moves around a properly-calibrated scale to indicate the value of the current flowing through the coil.

All such mechanisms respond to the magnitude of current through the coil and to the magnitude of the stationary magnetic field. The meter scale, however, may be calibrated to indicate a measurement of any quantity proportional to coil current.

AMMETERS

When direct currents that exceed the current-carrying capacity of the spring must be measured, the permanent-magnet and moving-coil mechanism is used in conjunction with a shunt. Figure 6 is a schematic representation of such an arrangement, where M represents the meter movement and R_S represents the shunt resistance. The instrument coil and shunt are connected in parallel; they divide the current in inverse ratio to their resistances. Instruments and shunts are calibrated as a unit and can be utilized to measure relatively-high direct currents. Shunts capable of carrying as much as 100 amperes frequently are assembled inside instrument cases, and external shunts are available in a variety of ratings up to many thousands of amperes.

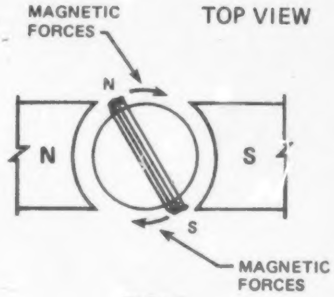


FIG. 5

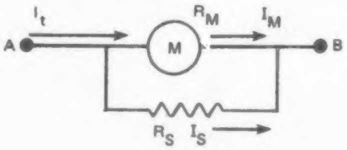


FIG. 6

VOLTMETERS

Stationary-magnet/moving-coil instruments also can be adapted to measure d.c. voltage. In this case, the instrument coil is connected in series with a resistor to limit the current to a value that safely can be carried by the coil and springs. Figure 7 shows the fundamentals of a d.c. voltmeter, where M is the meter movement and R_M is the series resistance.

Since the current is directly proportional to the applied voltage, the instrument can be calibrated in terms of d.c. volts. Series resistors, or "voltage multiplying resistors," as they sometimes are called, may be assembled within the instrument case or connected externally.

BASIC D.C. VOLTMETER CIRCUIT

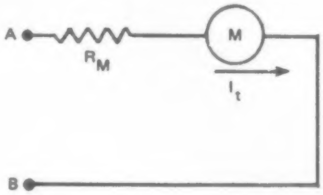


FIG. 7

OHMMETERS

A single-scale ohmmeter also can be constructed from a stationary-magnet/moving-coil meter movement (Figure 8). In this case, the instrument coil M is connected in series with a battery and a variable resistor R_y.

SINGLE SCALE OHMMETER CIRCUIT

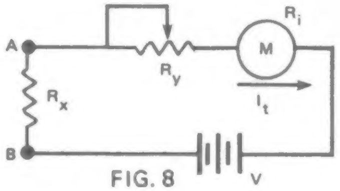


FIG. 8

The unknown resistor, R_x, placed across the ohmmeter at points A and B, becomes a part of the circuit and can be measured using simple d.c. theory. The scale of the ohmmeter, calibrated in ohms, indicates the resistance directly.

SUMMARY

Meter movements are energy convertors that change part of the energy of a magnetic field into mechanical potential energy stored in a coiled spring. A needle attached to the moving coil (or magnet) indicates the amount of energy stored in the spring and, thus, the magnitude of the current flowing through the coil of the meter. Appropriate circuitry can be used to calibrate the meter to measure voltages, currents, and unknown resistances.

DETERMINATION OF METER INTERNAL RESISTANCE

MATERIALS

Digital voltmeter (optional).
Volt-Ohm-Milliammeter (VOM).
2.5 K variable resistor.
Low-voltage power supply.
Meter movement. [For this and all subsequent labs, a meter movement with a full scale current of 1 ma is assumed (I_{FS} = 1 ma).]
Clip leads.

PROCEDURE

Before a voltmeter, an ohmmeter, or an ammeter can be constructed, the internal resistance of the basic meter movement must be known. If the internal resistance is not given on the instrument, it can be determined by the series-resistor-method (Figure 9).

1. Connect components as shown schematically in Figure 9.

2. Set R to provide maximum resistance. Set power supply for an output voltage of approximately one volt.

3. Adjust R until the meter indicates full scale. Increase power-supply voltage if necessary.

4. Disconnect the circuit and measure R with the VOM. Call this reading R_1; record in Data Table 1.

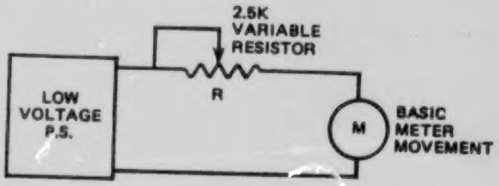


FIG. 9

5. Reconnect the circuit and adjust R until the meter indicates 1/2 scale. (Do not change power-supply voltage.) Measure R as before and call this reading R_2. Record.

6. The internal resistance of the meter R_M is determined from the following equation (see sample calculation):

$$R_M = R_2 - 2 R_1.$$

Record the value of R_M in Data Table 1.

SAMPLE CALCULATION

$$R_1 = 1256 \text{ ohm}$$
$$R_2 = 2620 \text{ ohm}$$
$$2\,R_1 = 2512 \text{ ohm}$$
$$R_M = R_2 - 2\,R_1$$
$$R_M = 2620\ \Omega - 2512\ \Omega$$
$$R_M = 108\ \Omega$$

7. If R_M is small enough, it can be read directly with a digital voltmeter.

CONSTRUCTION OF A D.C. VOLTMETER

MATERIALS

VOM.
Multiplier resistor, R.
Basic meter movement ($I_{FS} = 1$ mA, $R_M \approx 100\ \Omega$).

PROCEDURE

1. Prior to completing the basic voltmeter circuit shown in Figure 10, determine maximum voltage V_{AB} that will be measured, the meter's internal resistance R_M, and the meter's full-scale current I_{FS}.

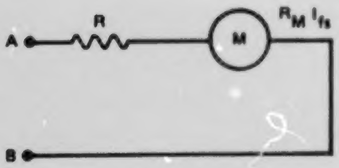


FIG. 10

2. Record the results of the following calculations in Data Table 2. In this circuit, the total resistance between A and B is —

$$R_t = R_M + R,$$

and the total current, the full scale meter current, is —

$$I_t = I_{FS}.$$

From Ohm's Law, the voltage V_{AB} is given by —

$$V_{AB} = I_t R + I_t R_M = I_{FS} R + I_{FS} R_M.$$

From this equation, the multiplier resistance is —

$$R = \frac{V_{AB}}{I_{FS}} - R_M.$$

SAMPLE CALCULATIONS

For $V_{AB} = 10$ volt full scale

$I_{FS} = 1$ mA

$R_M = 100\ \Omega$

$$R = \frac{10\ \text{V}}{1\ \text{mA}} - 100\ \Omega = 9900\ \Omega$$

For $V_{AB} = 100$ volt full scale

$I_{FS} = 1$ mA

$R_M = 100\ \Omega$

$$R = \frac{100\ \text{V}}{1\ \text{mA}} - 100\ \Omega = 99900\ \Omega$$

Determine necessary series resistance R, and construct meter circuit.

3. Measure voltage of an unknown source (maximum voltage V_{AB}) with meter constructed, and record in Data Table 2. The measured voltage is V_{AB} times the fraction of full scale deflection registered by the needle.

4. Measure the same voltage with the VOM, and record in Data Table 2.

5. Explain voltmeter operation.

CONSTRUCTION OF A D.C. AMMETER

MATERIALS

VOM.
Shunt resistor, R_S (resistance wire may be used).
Basic meter movement ($I_{FS} = 1$ mA, $R_i = 100\ \Omega$).

PROCEDURE

1. Determine maximum total current I_t to be measured.

2. From basic ammeter circuit depicted in Figure 11, the shunt resistance can be calculated as follows:

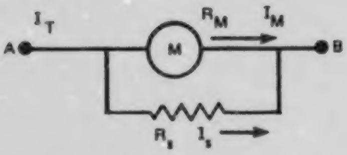


FIG. 11

V_M = Voltage across meter

V_s = Voltage across shunt

$$V_M = V_s.$$

$$V_M = I_M R_M$$

$$V_s = I_s R_s$$

$$I_t = I_s + I_M$$

$$R_s = \frac{R_M I_M}{I_t - I_M}$$

Record results in Data Table 3.

SAMPLE CALCULATION

Assume the maximum current to be measured is —

I_t = 100 mA (full-scale deflection)

R_M = 100 Ω

I_M = 1 mA

$$R_s = \frac{R_M I_M}{I_t - I_M}$$

$$R_s = \frac{(100\ \Omega)(1\ \text{mA})}{(100\ \text{mA} - 1\ \text{mA})}$$

$$R_s = 1.01\ \Omega$$

With such a shunt resistance in Figure 11, the meter movement would register full scale (100 mA), although only 1 mA passed through the meter-movement coil. The remaining 99 mA are carried by the shunt.

3. Determine and record necessary shunt resistance R_s, and construct meter circuit.

4. Measure unknown current with constructed meter (maximum current I_t) and record in Data Table 3. The unknown current is equal to I_t times the fractional full-scale deflection registered by the needle.

5. Measure the same current with the VOM and record in Data Table 3.

6. Explain any difference between current measured with constructed meter and that measured with VOM.

DATA TABLE 1: METER INTERNAL RESISTANCE

R_1 = ____________ Ω

R_2 = ____________ Ω

$R_M = R_2 - 2\,R_1 =$ ____________ Ω

DATA TABLE 2: D.C. VOLTMETER

V_{AB} = ____________ V

I_{FS} = ____________ A

R_M = ____________ Ω

$R = \dfrac{V_{AB}}{I_{FS}} - R_M =$ ____________ Ω

Unknown voltage measured with constructed meter: ____________ V

Same voltage measured with VOM: ____________ V

DATA TABLE 3: D.C. AMMETER

I_t = ____________ mA

I_M = ____________ mA

R_M = ____________ Ω

$R_s = \dfrac{R_M I_M}{I_t - I_M} =$ ____________ Ω

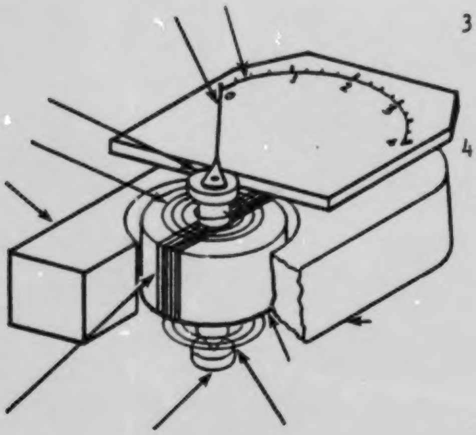
Unknown current measured with constructed meter: ____________ mA

Same current measured with VOM: ____________ mA

SELF CHECK

1. List three types of meter movements and an application of each.

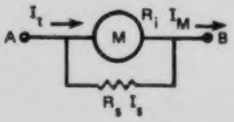
2. Identify the components of the adjoining meter movement.



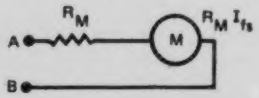
3. Explain how a stationary-magnet/moving-coil meter operates. Include a discussion of the meter movement as an energy convertor.

4. Correctly identify the following meter circuits.

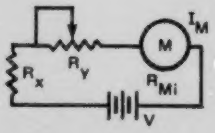
   a.



   b.



   c.



5. Explain how the internal resistance of a meter movement can be determined.

6. In a stationary-magnet/moving-coil ammeter the measured current is proportional to the coil current. If the measured current is 0.1 amps, where the coil current is 0.0001 amps, what is the measured current when the coil current is 0.0006 amps?

7. In a circuit like the one shown in Figure 7, find the voltage between points A and B if $I_t = 0.001$ A, $R_m = 10{,}000\Omega$ and the voltage drop across the meter is 0.1 volts.

8. Consider a circuit like the one shown in Figure 8. Let $V = 12$ volts, $R_y = 1000\Omega$ and the voltage drop across the meter be negligible. If $I_t = 10$ mA, what is the value of R_x?

9. An ohmmeter is designed to have a linear scale. If full scale deflection indicates a resistance of $20{,}000\Omega$, what resistance does a half scale deflection indicate?

10. If the ohmmeter of Problem 9 is designed to have a logarithmic scale, what resistance does half-scale deflection correspond to?

10T2

BIMETALLIC STRIPS

A bimetallic strip is constructed to two dissimilar metals having different coefficients of linear expansion and bonded together in intimate contact along their long dimension. When the strips are heated, the metal with the higher coefficient of expansion expands more than the other metal, causing the strip to bend. Because the amount of bending is proportional to the temperature change, bimetallic strips are convenient devices for use in thermostats and oven thermometers. The bimetallic strip can be considered a transducer, since the thermal energy absorbed by the strip is converted partially into the potential energy of deformation stored in the bent strip.

This module introduces the design and operation of bimetallic strips. In the laboratory, the student will construct a simple bimetallic strip; will measure the motion of the end of the strip as it is heated; and will use it to construct a control circuit.

OBJECTIVES

Upon completion of this module, the student should be able to:

- Describe how a bimetallic strip operates.
- List at least two applications of bimetallic strips.
- Name the combination of metals often used in bimetallic strips.
- Given the length, thickness, temperature, and deflection of a straight bimetallic strip, calculate the required difference in coefficients of expansion of the two metals.
- Given the radius, thickness, length, coefficient, and distance of deflection of a spiral bimetallic strip, calculate the temperature increase.

- Given the appropriate equipment, construct a simple bimetallic strip, measure its deflection when heated, and assemble a control circuit.

DISCUSSION

Bimetallic strips are constructed of two different metals fastened together along their length. The construction is indicated in Figure 1. Figure 1a shows a piece of bimetallic strip made of two equal lengths of two different metals. The two metals exhibit different values for their coefficients of linear expansion. Assume that metal A has a higher value for its coefficient of linear thermal expansion than metal B. The strip is formed at room temperature by bonding or fastening the two pieces of metal together.

In order to manufacture bimetallic strips in large quantity, strips of two separate metals are fed into a welding station and spot welded together along their length. At room temperature the strip is, of course, straight.

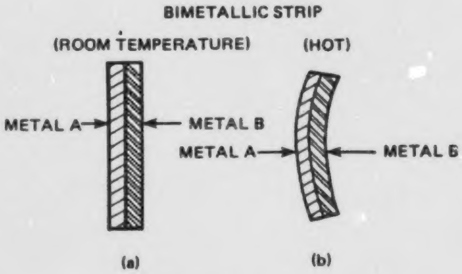


FIG. 1

Now, what happens when the strip is heated? The answer is found in Figure 1b. Both metals expand as the temperature rises, but metal A expands more. Since the two metals are bonded together, and since metal A cannot simply slip past metal B by expanding, a force is generated and bends the strip as shown in Figure 1b. The subsequent bending allows metal A to become slightly longer, while remaining attached to metal B along its length. Thermally-generated energy and forces of expansion are responsible for the bending.

The two metals used in a bimetallic strip should have values of coefficient of linear expansion (Table 1) that differ as much as possible; for if the two metals had equal values of the coefficient, they would expand equally, and there would be no bending. The amount of bending increases as the difference between the coefficients increases.

TABLE 1. PROPERTIES OF VARIOUS MATERIALS

Material	Linear Coefficient of Thermal Expansion (Per Degree C)	Modulus of Elasticity (Dyne/cm^2)
Aluminum (commercially pure)	23.5×10^{-6}	6.9×10^{11}
Brass (yellow, 65% Cu, 35% Zn)	20.3×10^{-6}	10.3×10^{11}

Copper (commercially pure)	16.5×10^{-6}	12.8×10^{11}
Glass		
Borosilicate	6.2×10^{-6}	---
Pyrex	$(3.1 - 3.5) \times 10^{-6}$	---
Soda-lime	$(7.5 - 9.2) \times 10^{-6}$	---
Invar (64% Fe, 36% Ni	1.2×10^{-6}	14.8×10^{11}
Kovar (54% Fe, 28% Ni, 18% Co)	5.9×10^{-6}	13.8×10^{11}
Quartz, fused	0.5×10^{-6}	7.2×10^{11}
Stainless steel (17-20% Cr, 8-11% Ni)	9.6×10^{-6}	19.3×10^{11}
Titanium (commercially pure)	8.5×10^{-6}	11.6×10^{11}
Tungsten (commercially pure)	4.3×10^{-6}	34.5×10^{11}

The largest value of the difference can be obtained if one metal has a high value for the coefficient of linear expansion and the other metal has a much lower value. The metal "invar" has a relatively low value for its coefficient. Invar is an alloy, consisting of 64% iron and 36% nickel. Its coefficient of linear expansion is about 1.2×10^{-6} per C°, a value much lower than that of most other metals. Thus, invar often is used as the metal with low expansion characteristics.

Brass (65% copper, 35% zinc) is a metal with a reasonably high value of thermal expansion coefficient (about 20.3×10^{-6} C°), about 20 times greater than invar. Although this is not the highest known value, brass is relatively inexpensive. Thus, the combination of invar-brass offers a relatively high value for the difference in thermal expansion coefficients; this combination often is used in practical bimetallic strips.

Choice of metals depends upon the temperature range in which the device is to be used. The combination of brass and invar frequently is used for applications in which the expected temperature range centers around room temperature. The coefficients of thermal expansion change with temperature, however, so that no one combination of metals is suitable for all temperature ranges. For measurements in high temperature ranges, alloys consisting of 20-24% nickel, 2-4% chromium, and the remainder of iron have been useful.

With proper choice of metals, bimetallic devices have been constructed to cover a temperature range from -300°F up to 1000°F.

Construction of bimetallic strips can take a variety of forms, as seen in the straight strip (Figure 1). Although the operation of the device is understood readily with reference to the straight strip, other configurations probably are used more often in practical devices. Spiral construction (Figure 2) often is used in practical temperature-measuring systems, as the spiral will either unwind or coil more tightly as the temperature changes, which causes the motion of the end of the spiral. The deflection of the end of the spiral is denoted as x in Figure 2.

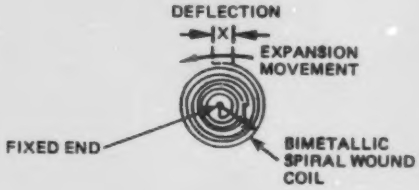


FIG. 2

The stress pattern generated in the bimetallic strip as it is heated is shown in Figure 3. Metal with the higher coefficient of linear expansion is the bottom one. Ordinarily, when it is heated, it will expand according to Equation 1.

$$\ell_2 - \ell_1 = \alpha \ell_1 (T_2 - T_1) \qquad (1)$$

where

ℓ_1 = length of metal sample at some lower temperature.
ℓ_2 = length of metal sample at some higher temperature.
α = coefficient of linear expansion for sample.
T_1 = lower temperature.
T_2 = higher temperature.

Equation 1 gives the maximum elongation of an unbonded piece of metal. However, when a piece of this same metal is bonded to some other metal having a lower expansion coefficient, it is constrained from expanding linearly as much as the above relationship would require. This constraint leads to compressive forces being generated in that metal at the interface between the two metals. Similarly, the low-expansion metal at the top is being stretched by its contact with the high-expansion metal. This process leads to tension (i.e., forces that tend to stretch the material) at the interface. The combination of these two forces causes the strip to bend, so that it is concave upwards. After bending, the top of the strip becomes slightly compressed and the bottom slightly stretched, leading to the pattern of compression and tensile forces illustrated in Figure 3.

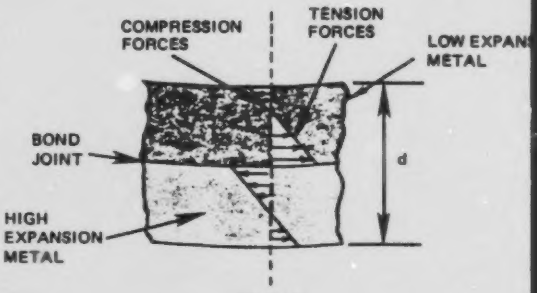


FIG. 3

Consider the deflection that is experienced by a bimetallic strip clamped at one end. The geometry is illustrated in Figure 4; length of strip is denoted by ℓ and its thickness by d. The strip is assumed to be straight at some value of temperature T_1. When the temperature is raised to T_2, the strip will bend (Figure 4); and the end of the strip will move by an amount x. Deflection is given by Equation 2.

$$x = 3\,(\alpha_A - \alpha_B)(T_2 - T_1)\ell^2/4d \qquad (2)$$

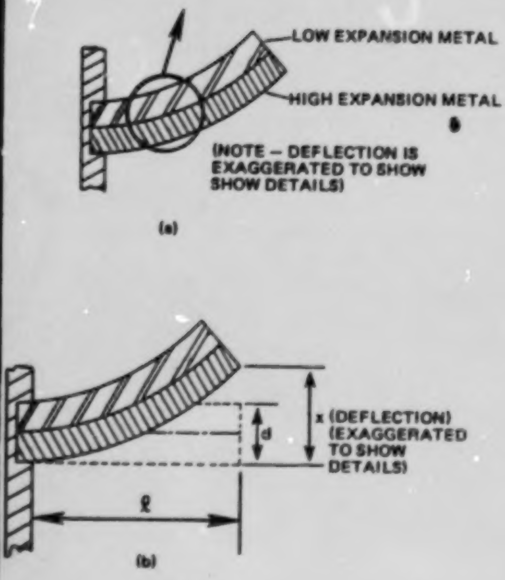

LOW EXPANSION METAL
HIGH EXPANSION METAL
(NOTE – DEFLECTION IS EXAGGERATED TO SHOW SHOW DETAILS)
(a)
x (DEFLECTION) (EXAGGERATED TO SHOW DETAILS)
d
ℓ
(b)

FIG. 4

where

x = deflection of end of strip.
α_A = coefficient of thermal expansion of metal A.
α_B = coefficient of thermal expansion of metal B.
T_1 = lower temperature at which strip is straight.
T_2 = higher temperature, which produces amount of deflection x.
ℓ = length of strip.
d = thickness of strip.

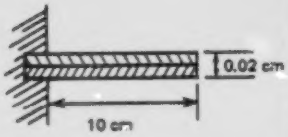
Equation 2 is employed in the solving of a problem in Example A.

The deflection of a straight strip, therefore, is directly proportional to the change in temperature and to the square of the length of the strip, and inversely proportional to its thickness. The fact that the deflection is proportional to the temperature change is a valuable feature for the practical use of bimetallic strips, as this simplifies its calibration. The device can be calibrated at two known temperatures, and a linear response over the entire temperature scale is obtained, as long as α_A and α_B vary negligibly with temperature over this range.

In practice, the linear response of the bimetallic strip extends over a limited temperature range. For an application involving measurements over a wide range of temperatures, make sure that the materials of the chosen device exhibit linearity over the desired range of temperatures (Example A).

EXAMPLE A: BIMETALLIC STRIP DEFLECTION

Given: A straight bimetallic strip of brass and Invar is clamped at one end and is 10 cm long and 0.02 cm thick.



Find: Deflection of end when strip temperature increases by 10C°.

Solution: Refer to Table 1: For brass, the linear coefficient of thermal expansion is 20.3×10^{-6} per C° and for Invar it is 1.2×10^{-6} per C°.

Use Equation 2:

$$x = 3(\alpha_A - \alpha_B)(T_2 - T_1)\ell^2/4d$$

$$= 3(20.3 \times 10^{-6} - 1.2 \times 10^{-6})(10^2)10/(4 \times 0.02)$$

$$x = 0.72 \text{ cm}$$

For a spiral (Figure 2), the deflection of the end of the spiral is given by Equation 3.

$$x = 9\ (\alpha_A - \alpha_B)(T_2 - T_1)r\ell/4d \qquad (3)$$

where

r = radius of spiral.
ℓ = extended length of strip.

Equation 3 is valid only under conditions in which the two metals have nearly equal moduli of elasticity (Table 1). Thus, in the case of a spiral, the deflection of the end again is linearly proportional to the temperature change.

An important use of bimetal strips is in thermostats. Bimetal strips are found in many thermostats used for control of temperature in the home. If the thermostat cover is removed, the strip (often a spiral) can be viewed. As the temperature in the home changes, the motion of the strip is used to open and to close a switch, which turns the heating/cooling system on and off. Thus, the bimetallic strip is an important element in a common house control system.

Oven thermometers also use bimetallic strips and usually have a round face, to accommodate a spiral strip. Motion of the bimetallic strip moves a pointer that indicates the oven temperature.

SUMMARY

The bimetallic strip is a good example of a practical application of the expansion of materials. It consists of two strips of metal fastened together in such a manner that at some temperature they are straight. The two strips of metal have different coefficients of expansion. If the strip is heated, it will bend in a curve; the amount of bending produced will be determined by the change in temperature experienced by the strip. Bimetallic strips, therefore, are frequently used in thermostats.

BIMETALLIC STRIP DEMONSTRATION

MATERIALS

Brass and zinc strips (other metal strips can be used, as available).
Metal snips.
Pliers.
Vise.
Piece of thin bell wire.
Ruler.
Candle and matches.
Oven thermometer.
Screwdriver.
Hammer, nail, and tacks.
Board (6" x 6" x 1/2").
Battery (1 1/2 volt).
Bulb (1 1/2 volt) and holder.
Wire and clips as needed.
Hair dryer or substitute.
Epoxy.

PROCEDURE

1. Construct a simple bimetallic strip as follows:

   (a) Cut off equal lengths of brass and zinc strips, about 12 inches each. Smooth to straighten. The brass and zinc are suggested because of easy availability; strips of other metals can be substituted as available.

   (b) Clamp the two strips together, and make slits in strips with metal snips (Figure 5). Bend flaps over to fasten two strips together. Tighten flaps with pliers to hold securely. Make as many flaps as possible.

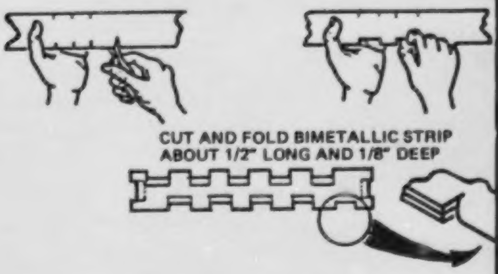

CUT AND FOLD BIMETALLIC STRIP
ABOUT 1/2" LONG AND 1/8" DEEP

FIG. 5

2. Observe the deflection of strip as it is heated.

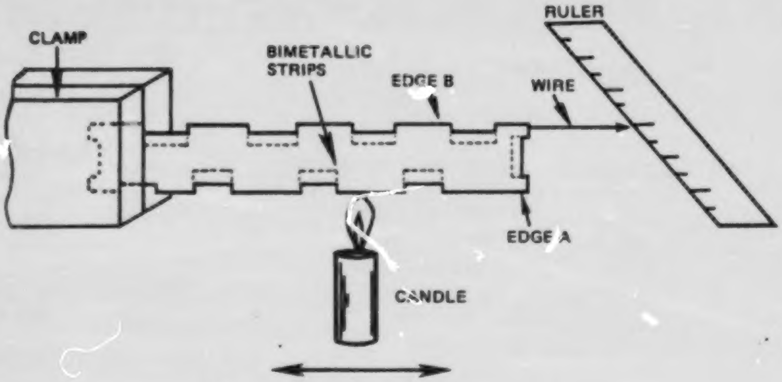


FIG. 6

(a) Set up apparatus (Figure 6). Firmly clamp one end of strip. Insert a piece of thin wire into flap at free end of strip to be used as a pointer. Place a ruler at right angles to strip with end of wire near ruler.
(b) Light candle and move flame back and forth under strip on edge A.
(c) Observe motion of strip.
(d) Remove flame. Does wire return to its original position?
(e) Reverse strip edge A for edge B, and repeat steps b and c. Which direction does pointer move now? Why?

3. Disassemble an inexpensive oven thermometer, and use the bimetallic spiral strip for control of a light bulb.

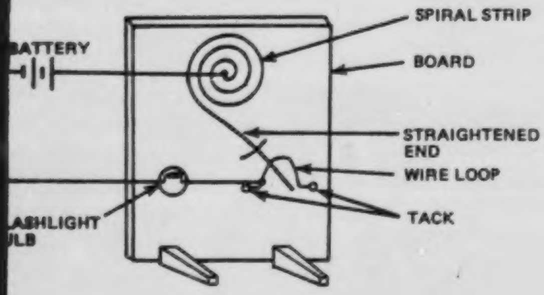


FIG. 7

(a) Pry apart an inexpensive oven thermometer with a screwdriver. Remove spiral bimetallic strip.
(b) Straighten end of bimetallic strip (about last three inches). Set up apparatus shown (Figure 7). Make a nail hole in board, and firmly insert spindle from strip into hole and epoxy it to board so that it does not turn. Mount a wire loop with tacks with straightened end of spiral through loop. Strip should be about 1/8" from loop. Make electrical connections as shown, with connections from battery to spindle of spiral and from the bulb to wire loop.
(c) Heat spiral with hair dryer. Does bulb light? (If it does not light quickly, reduce gap between loop and end of strip.)
(d) Remove hair dryer. Does light bulb go out? How does this experiment demonstrate use of a bimetallic strip for control applications?

SELF CHECK

1. Describe operation of bimetallic strips.

2. Name at least two applications for bimetallic strips.

3. Name a combination of metals often used in bimetallic strips.

4. If a straight bimetallic strip of length 18 cm and of thickness 0.1 cm is clamped at one end and heated by 15C°, what is the required difference in the coefficients of expansion of metals, for an observed motion of free end of 0.23 cm?

5. A spiral bimetallic strip of radius 5 cm, thickness 0.05 cm and length 25 cm, is made of metals with coefficients of thermal expansion equal to 17×10^{-6} and 5×10^{-6} per C°. When it is heated, the end deflects through a distance 0.675 cm. What was the increase of temperature?

6. What is the coefficient of linear expansion of a metal whose length changes from 50 to 50.1 cm when the metal is heated from 10°C to 100°C?

7. A certain steel strip has a coefficient of linear expansion of 11×10^{-6}/C°. What should the temperature of the rod be in order that the percent increase in the rod's length be 1% if the rod is initially at a temperature of 23°C?

8. What should the thickness of a bimetallic strip be if the strip is to be 8 cm long and have a deflection of 0.5 cm when the temperature of the strip is raised by 60 C°? The strip is made out of copper and tungsten.

10T3

TEMPERATURE MEASURING DEVICES

Temperature is a measure of the average kinetic energy of the molecules of a substance. The increased motions of molecules caused by the absorption of heat energy is difficult to measure directly, but these microscopic motions cause changes in other physical quantities that can be measured. Thus, measurements of physical quantities like volume, pressure, mechanical distortion, voltage, and resistance can be used an an indication of temperature. Ultimately, temperature in a particular range must be defined in terms of the measurement of such a quantity. A temperature-measuring device can be viewed as one which transforms molecular-kinetic energy into a form that is more easily accessible to the experimenter. This module introduces the theory and applications of such devices. In the laboratory, the student will perform temperature measurements with a thermocouple and with a thermistor.

OBJECTIVES

Upon completion of this module, the student should be able to:

- Sketch a fluid-filled thermometric system, label the pointer, pressure-sensitive element, scale, capillary tube, and fluid-filled bulb; and describe in one or two sentences the friction of each.
- Describe in a short paragraph the theory and operation of a bimetallic strip, and indicate how it can be used in a practical situation.
- Draw, label, and briefly describe the operation of the five key parts of a thermocouple.
- Given one of the three variables of a thermocouple, including

its abbreviation, symbol, or color code, identify the two conductors and the other two variables.

- Describe the use of the thermopile in a temperature-sensing system.
- Discuss in two or three sentences why and how a thermistor is used to measure small temperature variations.
- Given the appropriate equipment, measure the temperature of a liquid with a thermocouple and with a thermistor, and plot the temperature measured against their output voltage.

DISCUSSION

FLUID-FILLED DEVICES

A common temperature-measurement technique makes use of fluid-thermal systems. When a closed system is filled with a fluid and sealed, an increase in temperature causes the fluid to expand. When the fluid expands, the motion is transmitted to a pointer or indicator by means of a link and lever system.

Fluid-filled thermal systems are categorized by the type of fluid used. The fluid may be a liquid or a gas, or in the case of mercury, a fluid metal.

All such measuring devices consist of a bulb (reservoir for the fluid), a pressure-sensitive element and a capillary tube joining the two. An example of such a device is shown in Figure 1.

Liquid and mercury systems typically operate on a volumetric change of the filled material with a change in temperature, while gas systems operate on a pressure change with temperature.

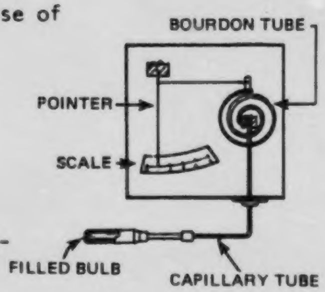


FIG. 1

The most common type of fluid system is the glass thermometer, which uses either mercury of alcohol as the fluid. Temperature changes cause a change in the volume of the fluid, and the movement of the liquid along a thin cylinder of constant diameter permits the change in volume to be registered as a change in length. Thus, an empirically-constructed linear scale along the cylinder can be calibrated easily in terms of temperature.

More sophisticated systems employ a transducer that converts the movement caused by volumetric or pressure changes to a scaled pneumatic or electronic signal. This signal, in turn, is transmitted to a receiving instrument, where the signal response is interpreted as a temperature. These temperature transmitters are used in industry for control purposes; on motor vehicles for measurement of engine oil temperature; and also as temperature control devices in the home.

BIMETALLIC STRIPS

A bimetallic strip is constructed by the joining together of two metals that have different rates of thermal expansion. When the device is heated, the unequal thermal expansion of the two metals causes a bending of the strip, because one metal expands at a greater rate than the other. Figure 2 shows how a temperature change is transformed to mechanical motion in the bimetallic strip. Bimetallic strips of this construction are used most commonly in single-temperature control circuits to

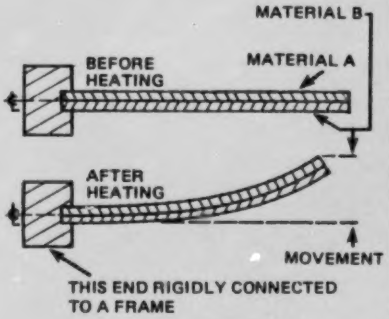


FIG. 2

make or to break an electrical contact that produces corresponding fluctuations in temperature. Figure 3 illustrates one such device. When the temperature of the unit increases to a point at which the thermal expansion of the bimetallic strip is great enough to overcome the spring tension, the contact opens, breaking the electrical circuit to the heater. As the unit cools down, the bimetallic strip returns to its original shape and again closes the heater circuit. The cycle continues to repeat and, therefore, to control temperature in a given enclosure.

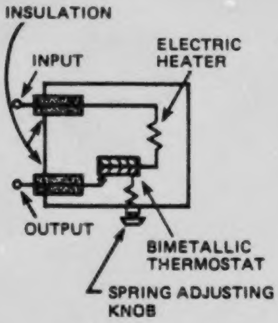
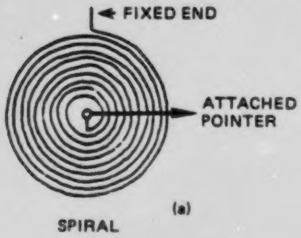
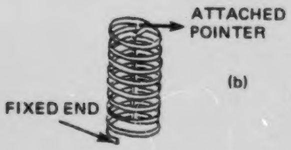


FIG. 3

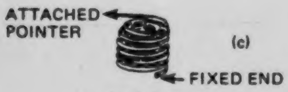
In applications where it is desirable to incorporate the response of a bimetallic strip to provide temperature measurement instead of control, the bimetal element is wound into the form of a spiral or helix (Figure 4). The response of movement of the strip with temperature change is increased in this way and can be made linear over a range of temperature change of about 400°F. For temperatures greater than 400°F, nonlinear scales can be empirically constructed.



SPIRAL



SINGLE HELIX



MULTIPLE HELIX

FIG. 4

THERMOCOUPLES

Measurement of temperatures by their electrical effect is used often in applications where it is necessary to take measurements in remote or hazardous locations. When a thermocouple is used to measure temperature, the read-out device can be in a central control room thousands of feet or more from the point of measurement. The voltage produced by the temperature differential sensed in the thermocouple can be transmitted easily over electrical cables.

Two dissimilar metals joined together at one end produce a small voltage across the open ends if a temperature difference exists between the junction and the open ends. The open ends are connected to a sensitive voltage-measuring device, such as a millivoltmeter or a potentiometer (Figure 5).

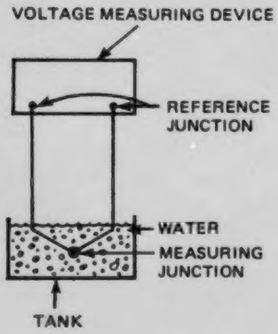


FIG. 5

The point at which the wires are connected is called the "measuring junction." The points at which the wires are connected to a voltage-measuring device are called "reference junctions." When the temperatures of the measuring and reference junctions are equal, no voltages will be developed at the voltmeter. When the temperature of the measuring junction is elevated above that of the reference junction, a voltage will be generated with some polarity (plus-or-minus). When the measuring-junction temperature is decreased below that of the reference junction, the voltage developed at the voltmeter reverses polarity. As the temperature difference between the two junctions increase, the voltage at the measuring junction increases.

Table 1 lists common thermocouple abbreviations, symbols, conductor polarities, and wire color codes; Table 2 provides thermocouple data for design use.

TABLE 1. THERMOCOUPLE SYMBOLS, MATERIALS, POLARITY, AND COLOR CODES

Symbol	Positive Conductor (+)			Negative Conductor (-)		
	Material	Abbreviation	Color Code	Material	Abbreviation	Color Code
E	Chromel	CR	Purple	Constantan	C	Red
J	Iron	Fe	White	Constantan	C	
T	Copper	Cu	Blue	Constantan	C	
K	Chromel	CR	Yellow	Alumel	AL	
R	Platinum with 13% Rhodium	Pt/Rh-13	Black	Platinum	Pt	
S	Platinum with 10% Rhodium	Pt/Rh-10		Platinum	Pt	

Thermocouple calibration tables relate millivolts generated by the thermocouple to the temperature generated by the thermocouple to the temperature difference between the two junctions. Such tables are compiled from test data obtained when the reference-junction temperature is fixed, when the measuring-junction temperature is varied, and when the voltage generated by the thermocouple is measured precisely. Calibration tables often are based on constant reference-junction temperatures of 32°F (0°C), 75°F or 25°C. With the help of these tables, a thermocouple, and a millivoltmeter, temperatures can be determined accurately. Example A illustrates the use of Table 2.

EXAMPLE A: SELECTION OF MEASURING-JUNCTION TEMPERATURE

Given:	An iron-constantan (type J) thermocouple has a reference-junction temperature of 32°F and a millivolt value of 3.708 mV.
Find:	Temperature of measuring junction.
Solution:	Notice that the reference junction temperature for Table 2 is 32°F. At 32°F, the voltage developed is 0.00. From the thermocouple calibration, Table 2, under the type J or iron-constantan, a mV value of 3.708 corresponds to 160°F, the temperature of the measuring junction.

TABLE 2. PARTIAL THERMOCOUPLE DATA

°F	mV Output for Thermocouple					°F	mV Output for Thermocouple				
	E	J	T	K	S		E	J	T	K	S
0	-1.026	-0.885	-0.674	-0.692	-0.092	110	2.627	2.233	1.752	1.748	0.256
10	-0.709	-0.611	-0.467	-0.478	-0.064	120	2.977	2.526	1.988	1.977	0.291
20	-0.389	-0.334	-0.256	-0.262	-0.035	130	3.329	2.820	2.226	2.206	0.328
30	-0.065	-0.056	-0.043	-0.044	-0.006	140	3.683	3.115	2.467	2.436	0.365
40	0.262	0.224	0.173	0.176	0.024	150	4.041	3.411	2.711	2.666	0.402
50	0.591	0.507	0.391	0.397	0.055	160	4.401	3.708	2.958	2.896	0.440
60	0.924	0.791	0.611	0.619	0.087	170	4.764	4.006	3.206	3.127	0.478
70	1.259	1.076	0.834	0.843	0.119	180	5.130	4.305	3.458	3.358	0.517
80	1.597	1.363	1.060	1.068	0.152	190	5.498	4.605	3.711	3.589	0.557
90	1.937	1.652	1.288	1.294	0.186	200	5.869	4.906	3.967	3.819	0.597
100	2.281	1.942	1.518	1.520	0.221	210	6.242	5.207	4.225	4.049	0.637

The voltage generated by a thermocouple is a function of the difference in temperature between the two junctions. A change in temperature of either junction causes a change in millivolt reading. The necessity arises, therefore, to ensure that the thermocouple output is caused only by a temperature variation at the measuring junction. This precaution can be observed when the reference-junction temperature is held at a constant value and when the millivolt value at that temperature is subtracted from the millivolt value read on the voltmeter. The difference in millivolts then is converted to temperature in the table. Example B illustrates this method. A practical procedure involves the use of a thermocouple table referenced to 32°F (0°C). The reference junction then may be placed in ice water, and the thermocouple output converted directly to temperature from the table (Figure 6). The millivolt value of the reference-junction temperature need not be subtracted, because it is equal to zero at the reference temperature.

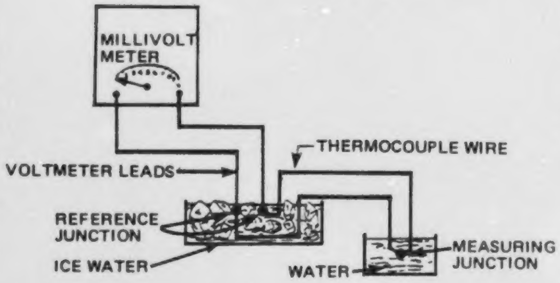


FIG. 6

EXAMPLE 8: CALCULATION OF THERMOCOUPLE OUTPUT

Given: For a chromel-constantan (type E) thermocouple, the measuring-junction temperature T_2 is 200°F and the reference-junction temperature T_1 is 100°F.

Find: Voltage generated by the thermocouple.

Solution: Remeber that the thermocouple output is a function of the temperature difference between the two junctions.

From calibration Table 2, under the type E column, 200°F corresponds to 5.869 mV, and 100°F to 2.281 mV. The difference is (5.869 - 2.281) = 3.588 mV. Thus, 6.00 mV is the thermocouple output.

THERMOPILE

When two or more thermocouples are connected in series, the voltage from the individual thermocouples add to produce a total voltage that is easier to measure and that can be used for control circuits. Figure 7 shows such an arrangement, giving a voltage five times as great as that measured with a single thermocouple. This arrangement is called a thermopile and can be used to measure small temperature differences.

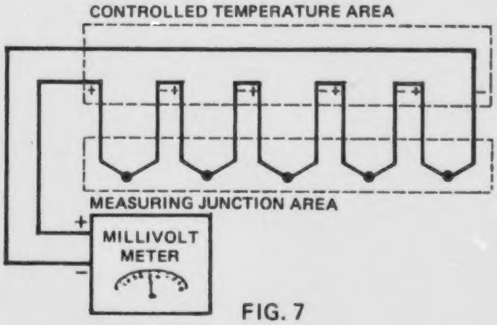

CONTROLLED TEMPERATURE AREA
+
-
MEASURING JUNCTION AREA
MILLIVOLT
METER

FIG. 7

THERMISTOR

Some electrical components change resistance with temperature to such a degree that measurement of resistance can be used as a means of determining very small temperature variations. Thermistors and elements called resistance-temperature devices are such components, which are used to measure temperature when connected to a bridge circuit. For example, the Wheatstone bridge circuit can be used to make accurate measurements of resistance (Figure 8).

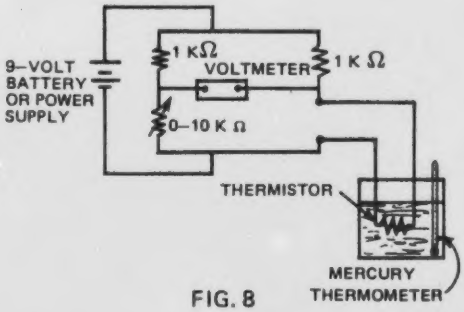

9—VOLT
BATTERY
OR POWER
SUPPLY
1 KΩ
VOLTMETER
1 K Ω
0–10 K Ω
THERMISTOR
MERCURY
THERMOMETER

FIG. 8

SUMMARY

Four methods of temperature measurements have been discussed:

- Fluid thermal systems.
- Bimetallic strip.
- Thermocouple and thermopile.
- Thermistor.

Fluid thermal and bimetallic strips are used in home thermostats for measurement and control and in fail-safe systems for gas-fired furnaces and hot water heaters.

Thermocouples are used in applications in which an electrical signal is convenient for temperature measurement and in which transmission of temperature information over a distance is desired. Thermistors and other resistance devices are used when high sensitivity is required.

THERMOCOUPLE TEMPERATURE MEASUREMENT

MATERIALS

Thermocouple wire (type J or copper-constantan).
Mercury thermometer.
Millivolt meter or strip-chart recorder.
Beaker.
Bunsen burner and stand
Graph paper.

Note: Thermocouple wire, thermocouple calibration tables, read-out devices, and other thermocouple application devices are available from various manufacturers. The National Bureau of Standards provides standard reference tables also.

PROCEDURE

This experiment involves the construction of a calibration curve for a thermocouple.

1. Construct a thermocouple by twisting two ends of thermocouple wire together and connecting other ends to a millivoltmeter. Place twisted end in a beaker of cold water (Figure 9). Measure water temperature with mercury thermometer, and record temperature and millivolt value in Data Table 1.

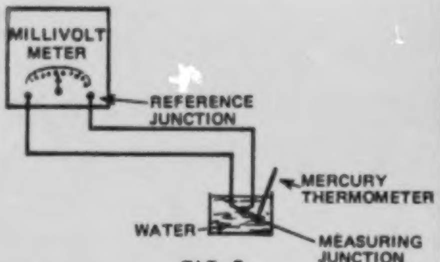


FIG. 9

2. Elevate temperature of water in approximately five-degree increments, and record each thermometer temperature and its corresponding millivolt reading in Data Table 1. Reference junction should be held constant, at room temperature.

3. Note condition when temperature of measuring junction and reference junctions are equal (that is, at room temperature). What is the reading on the millivoltmeter? What happens when the measuring junction temperature is higher than reference junction temperature? What happens when measuring junction temperature is lower than the reference junction temperature?

4. Plot temperature versus millivolt reading on linear graph paper.

Is the plot linear? Could it be used for temperature measurement?

THERMISTOR TEMPERATURE MEASUREMENT

MATERIALS

Two 1000 ohm resistors.
One 0-10,000 ohm variable resistor.
One 9-volt battery or power supply.
One voltohmmeter.
One thermistor with resistance of about 5000 ohms at 25°C, available from Yellow Spring Instrument Company, Yellow Springs, Ohio (or equal supplier).
One mercury thermometer.
Graph paper.
Crushed ice (1 cup).
Bunsen burner and stand.

PROCEDURE

The following procedure is designed to measure the temperature of a substance over a wide range using a thermistor, and to make a plot of temperature versus voltage.

1. Assemble the circuit shown (Figure 8).

2. With thermistor at room temperature, vary the variable resistor until bridge is balanced; that is, until voltmeter reads essentially zero volts. Do not change variable resistance throughout remainder of the experiment.

3. Change temperature in 5°F increments above and below ambient for about 100° range. This process can be accomplished by adding bits of ice (cooling) and with a bunsen burner (heating).

4. Using mercury thermometer as a standard, record in Data Table 2 temperature and corresponding voltage indicated by Wheatstone bridge at each step.

5. Plot temperature values against corresponding voltages on linear graph paper. This is a temperature-calibration curve. It can be used for precise temperature measurements.

6. Is the temperature calibration curve linear? Compare linearity,

accuracy, and sensitivity with the thermocouple method. How could the sensitivity of this temperature-measuring system be improved?

DATA TABLE 1

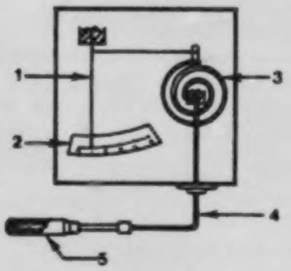
Temperature (In 5°F Increments)	Millivolt Value (Thermocouple Output)	Temperature (In 5° Increments)	Millivolt Value (Thermocouple Output)

DATA TABLE 2

Temperature (In 5°F Increments)	Millivolt Value (Thermocouple Output)	Temperature (In 5° Increments)	Millivolt Value (Thermocouple Output)

SELF CHECK

1. Label the parts of this fluid-filled thermal system and describe the functions of each part.



1 ______________________________

____

2 ______________________________

____

3 ______________________________

____

4 ______________________________

____

5 ______________________________

____

2. Briefly describe how a bimetallic strip can be used to control a thermostat.

3. Give a one- or two-sentence description of the function of the parts of a thermocouple, including the voltage measuring device, reference junctions, water, measuring junction, and tank.

4. Name the conductors, the abbreviation, and the color code for a thermocouple with a "T" symbol.

5. In one or two sentences, describe the use of a thermopile in a temperature-sensing system.

6. Describe in two or three sentences what happens to a thermistor when subjected to a temperature change.

7. An ideal gas has a temperature of 20°C when under a pressure of 1 atmosphere. If the volume of the gas is held fixed, what will be the temperature of the gas if the pressure is reduced to 0.9 atmospheres?

8. A chromel-alumel thermocouple has a reference temperature of 32°F. What is the temperature of the measuring junction when the voltage differential caused by the increase in temperature is 4.00 mV?

9. What voltage is generated when the measuring junction of a type J thermocouple is at a temperature of 210°F and the reference junction is at a temperature of 80°F?

10. At approximately what temperature should a type T thermocouple referenced to 75°F generate a voltage of -2.00 mV?

11. What should be the fractional increase in the volume of liquid in a thermometer when a thermometer is placed into a pan of boiling water if the original temperature of the liquid is 20°C and the coefficient of volume expansion for the liquid is 120×10^{-6}/C°.

10T4

THERMOCOUPLES

In 1826 Thomas Johann Seebeck discovered that a temperature difference that exists across a junction of two dissimilar metals is accompanied by a purely thermal emf. Since then, this type of junction, known as a "thermocouple," has become widely used as a thermometer.

Thermocouple theory and practical application are presented in this module. Examples are included to demonstrate the use of thermocouple calibration tables and computations necessary to determine temperature for nonstandard thermocouples. In the laboratory, the student will measure the thermoelectric voltage of different thermocouples.

OBJECTIVES

Upon completion of this module, the student should be able to:

- Sketch a typical thermocouple circuit, and briefly describe how it is used to measure temperature.
- Given the temperature of the reference junction, determine the following from calibration tables for a standard thermocouple:
  - a. Temperature of the test junction from the thermoelectric emf.
  - b. Thermoelectric emf from the temperature of the test junction.
- Calculate the thermoelectric emf of a nonstandard thermocouple.
- Determine the temperature of the test junction of a thermopile from the thermoelectric voltage.

320

- Calculate the thermoelectric power of a thermocouple, using the calibration tables.
- Given the appropriate equipment, measure the thermoelectric voltage of different types of thermocouples, and determine the temperature from calibration tables.

DISCUSSION

The thermocouple thermometer usually consists of two junctions of two dissimilar metals A and B, as illustrated in Figure 1. The reference junction R is maintained at a constant, known temperature T_R, and the test junction is the sensor for the unknown temperature T. The thermal emf V (also referred to as the "thermoelectric voltage" and the "thermoelectric emf") developed across the junctions is proportional to the difference in temperature $(T - T_R)$ between the two junctions. The sign of V depends upon which metal is connected to the millivoltmeter; that is, metal A or B. Its polarity changes when the temperature difference $(T - T_R)$ changes polarity.

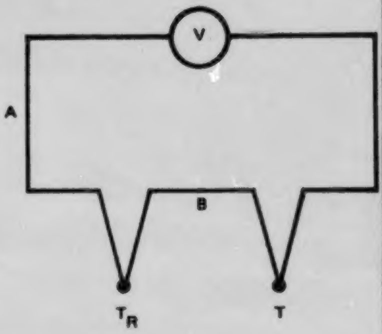


FIG. 1

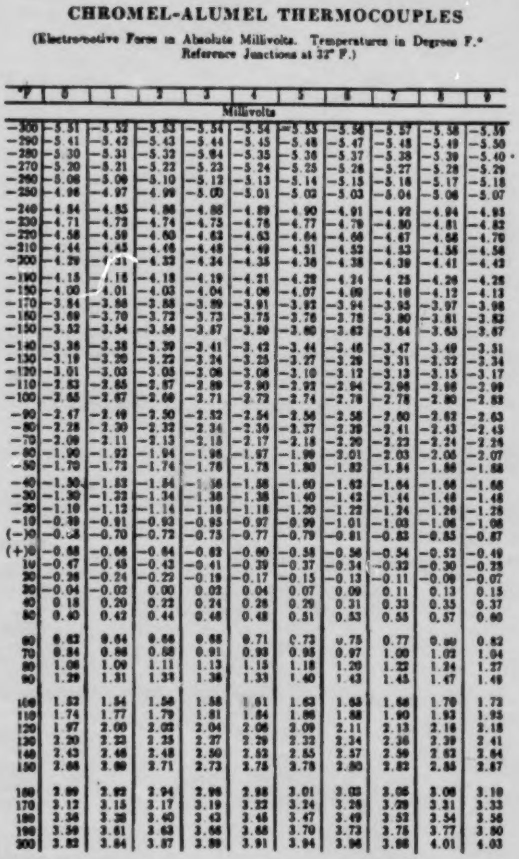
Calibration tables of thermoelectric voltage as a function of temperature T of the test junction referenced to $T_R = 0°C$ or 32°F are published for the standard thermocouples listed in Table 1 by type and components. Table 2 indicates the composition of alloys commonly used in thermocouples. Table 3 is a partial calibration table for a chromel-alumel thermocouple; Example A demonstrates its use.

TABLE 1. ACCEPTED THERMOCOUPLE DESIGNATIONS

TYPE	ALLOY COMPOSITION
J	Iron-Constantan
K	Chromel-Alumel
E	Chromel-Constantan
T	Copper-Constantan
S	Platinum vs Platinum - 10% Rhodium
R	Platinum vs Platinum - 13% Rhodium
B	Platinum - 6% Rhodium vs Platinum - 30% Rhodium

TABLE 2. COMPOSITION OF THERMOCOUPLE ALLOYS	
ALLOY NAME	COMPOSITION
Chromel	Nickel-Chromium
Alumel	Nickel-Aluminum
Constantan	Copper-Nickel

TABLE 3. SAMPLE CALIBRATION TABLE FOR TYPE K THERMOCOUPLES



CHROMEL-ALUMEL THERMOCOUPLES

(Electromotive Force in Absolute Millivolts. Temperatures in Degrees F.* Reference Junctions at 32° F.)

°F	0	1	2	3	4	5	6	7	8	9
	Millivolts									
−300	−5.51	−5.52	−5.53	−5.54	−5.54	−5.55	−5.56	−5.57	−5.58	−5.59
−290	−5.41	−5.42	−5.43	−5.44	−5.45	−5.46	−5.47	−5.48	−5.49	−5.50
−280	−5.30	−5.31	−5.32	−5.34	−5.35	−5.36	−5.37	−5.38	−5.39	−5.40
−270	−5.20	−5.21	−5.22	−5.23	−5.24	−5.25	−5.26	−5.27	−5.28	−5.29
−260	−5.08	−5.09	−5.10	−5.12	−5.13	−5.14	−5.15	−5.16	−5.17	−5.18
−250	−4.96	−4.97	−4.99	−5.00	−5.01	−5.02	−5.03	−5.04	−5.06	−5.07
−240	−4.84	−4.85	−4.86	−4.88	−4.89	−4.90	−4.91	−4.92	−4.94	−4.95
−230	−4.71	−4.73	−4.74	−4.75	−4.76	−4.77	−4.79	−4.80	−4.81	−4.82
−220	−4.58	−4.59	−4.60	−4.62	−4.63	−4.64	−4.66	−4.67	−4.68	−4.70
−210	−4.44	−4.45	−4.46	−4.48	−4.49	−4.51	−4.52	−4.53	−4.55	−4.56
−200	−4.29	−4.31	−4.32	−4.34	−4.35	−4.36	−4.38	−4.39	−4.41	−4.42
−190	−4.15	−4.16	−4.18	−4.19	−4.21	−4.22	−4.24	−4.25	−4.26	−4.28
−180	−4.00	−4.01	−4.03	−4.04	−4.06	−4.07	−4.09	−4.10	−4.12	−4.13
−170	−3.84	−3.86	−3.88	−3.89	−3.91	−3.92	−3.94	−3.95	−3.97	−3.98
−160	−3.69	−3.70	−3.72	−3.73	−3.75	−3.76	−3.78	−3.80	−3.81	−3.83
−150	−3.52	−3.54	−3.56	−3.57	−3.59	−3.60	−3.62	−3.64	−3.65	−3.67
−140	−3.36	−3.38	−3.39	−3.41	−3.42	−3.44	−3.46	−3.47	−3.49	−3.51
−130	−3.19	−3.20	−3.22	−3.24	−3.25	−3.27	−3.29	−3.31	−3.32	−3.34
−120	−3.01	−3.03	−3.05	−3.06	−3.08	−3.10	−3.12	−3.13	−3.15	−3.17
−110	−2.83	−2.85	−2.87	−2.89	−2.90	−2.92	−2.94	−2.96	−2.98	−2.99
−100	−2.65	−2.67	−2.69	−2.71	−2.72	−2.74	−2.76	−2.78	−2.80	−2.82
−90	−2.47	−2.49	−2.50	−2.52	−2.54	−2.56	−2.58	−2.60	−2.62	−2.63
−80	−2.28	−2.30	−2.32	−2.34	−2.36	−2.37	−2.39	−2.41	−2.43	−2.45
−70	−2.09	−2.11	−2.13	−2.15	−2.17	−2.18	−2.20	−2.22	−2.24	−2.26
−60	−1.90	−1.92	−1.94	−1.96	−1.97	−1.99	−2.01	−2.03	−2.05	−2.07
−50	−1.70	−1.72	−1.74	−1.76	−1.78	−1.80	−1.82	−1.84	−1.86	−1.88
−40	−1.50	−1.52	−1.54	−1.56	−1.58	−1.60	−1.62	−1.64	−1.66	−1.68
−30	−1.30	−1.32	−1.34	−1.36	−1.38	−1.40	−1.42	−1.44	−1.46	−1.48
−20	−1.10	−1.12	−1.14	−1.16	−1.18	−1.20	−1.22	−1.24	−1.26	−1.28
−10	−0.89	−0.91	−0.93	−0.95	−0.97	−0.99	−1.01	−1.03	−1.05	−1.07
(−)0	−0.68	−0.70	−0.72	−0.75	−0.77	−0.79	−0.81	−0.83	−0.85	−0.87
(+)0	−0.68	−0.66	−0.64	−0.62	−0.60	−0.58	−0.56	−0.54	−0.52	−0.49
10	−0.47	−0.45	−0.43	−0.41	−0.39	−0.37	−0.34	−0.32	−0.30	−0.28
20	−0.26	−0.24	−0.22	−0.19	−0.17	−0.15	−0.13	−0.11	−0.09	−0.07
30	−0.04	−0.02	0.00	0.02	0.04	0.07	0.09	0.11	0.13	0.15
40	0.18	0.20	0.22	0.24	0.26	0.29	0.31	0.33	0.35	0.37
50	0.40	0.42	0.44	0.46	0.48	0.51	0.53	0.55	0.57	0.60
60	0.62	0.64	0.66	0.68	0.71	0.73	0.75	0.77	0.80	0.82
70	0.84	0.86	0.88	0.91	0.93	0.95	0.97	1.00	1.02	1.04
80	1.06	1.09	1.11	1.13	1.15	1.18	1.20	1.22	1.24	1.27
90	1.29	1.31	1.33	1.36	1.38	1.40	1.43	1.45	1.47	1.49
100	1.52	1.54	1.56	1.58	1.61	1.63	1.65	1.68	1.70	1.72
110	1.74	1.77	1.79	1.81	1.84	1.86	1.88	1.90	1.93	1.95
120	1.97	2.00	2.02	2.04	2.06	2.09	2.11	2.13	2.16	2.18
130	2.20	2.23	2.25	2.27	2.29	2.32	2.34	2.36	2.39	2.41
140	2.43	2.46	2.48	2.50	2.52	2.55	2.57	2.59	2.62	2.64
150	2.66	2.69	2.71	2.73	2.75	2.78	2.80	2.82	2.85	2.87
160	2.89	2.92	2.94	2.96	2.98	3.01	3.03	3.06	3.08	3.10
170	3.12	3.15	3.17	3.19	3.22	3.24	3.26	3.29	3.31	3.33
180	3.36	3.38	3.40	3.43	3.45	3.47	3.49	3.52	3.54	3.56
190	3.59	3.61	3.63	3.66	3.68	3.70	3.73	3.75	3.77	3.80
200	3.82	3.84	3.87	3.89	3.91	3.94	3.96	3.98	4.01	4.03

210	4.06	4.08	4.10	4.12	4.15	4.17	4.19	4.21	4.24	4.26
220	4.28	4.31	4.33	4.35	4.38	4.40	4.42	4.44	4.47	4.49
230	4.51	4.54	4.56	4.58	4.61	4.63	4.65	4.67	4.70	4.72
240	4.74	4.77	4.79	4.81	4.83	4.86	4.88	4.90	4.93	4.95
250	4.97	4.99	5.02	5.04	5.06	5.08	5.11	5.13	5.15	5.17
260	5.20	5.22	5.24	5.26	5.29	5.31	5.33	5.35	5.38	5.40
270	5.42	5.44	5.47	5.49	5.51	5.53	5.56	5.58	5.60	5.62
280	5.65	5.67	5.69	5.71	5.73	5.76	5.78	5.80	5.82	5.85
290	5.87	5.89	5.91	5.94	5.96	5.98	6.00	6.02	6.05	6.07
300	6.09	6.11	6.13	6.16	6.18	6.20	6.22	6.25	6.27	6.29
310	6.31	6.33	6.36	6.38	6.40	6.42	6.45	6.47	6.49	6.51
320	6.53	6.56	6.58	6.60	6.62	6.65	6.67	6.69	6.71	6.73
330	6.76	6.78	6.80	6.82	6.84	6.87	6.89	6.91	6.93	6.96
340	6.98	7.00	7.02	7.04	7.07	7.09	7.11	7.13	7.16	7.18
350	7.20	7.22	7.24	7.26	7.29	7.31	7.33	7.35	7.38	7.40

EXAMPLE A: USE OF THERMOCOUPLE CALIBRATION TABLES

Given: A chromel-alumel thermocouple has $T_R = 32°F$ and $V = 2.73$ millivolts.

Find: The temperature of the test junction.

Solution: The temperature in increments of 10°F is listed in the first column of Table 3. Succeeding columns list the thermoelectric voltage as the temperature is increased in steps of 1°F. The emf of 2.73 mV is located in the row labeled 150°F and in the column labeled 3°F. Therefore —

$$T = 150°F + 3°F$$

$$T = 153°F$$

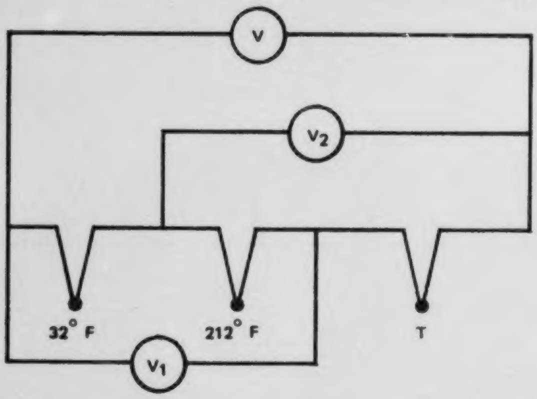
If the temperature of the reference junction is not 32°F, the temperature of the test junction still can be found from the tables referenced to 32°F. Examples B and C illustrate this procedure.

EXAMPLE B: ARBITRARY REFERENCE TEMPERATURE

Given: A chromel-alumel thermocouple has $T_R = 212°F$ and $V = 3.28$ mV.

Find: Temperature T of the test junction.

Solution: The thermocouple circuit can be represented by three junctions, at 32°F, at 212°F, and at the unknown temperature T, all in series. The thermoelectric voltage is the sum of the voltage V_1 between 32°F and 212°F and the voltage V_2 between 212°F and T.



From Table 3, $V_1 = 4.10$ mV; therefore, the sum V is

$$V = V_1 + V_2$$
$$= 4.10 \text{ mV} + 3.28 \text{ mV}$$
$$V = 7.38 \text{ mV}$$

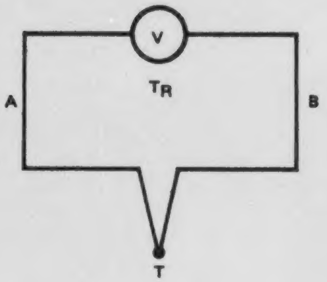
According to Table 3, the temperature corresponding to V = 7.38 mV is T = 358°F.

EXAMPLE C: SINGLE JUNCTION THERMOCOUPLE

Given: The thermoelectric voltage across a single-junction thermocouple is 2.80 mV. Room temperature is 72°F.

Find: The temperature of the test junction.

Solution: The reference temperature for the single junction is the temperature at the meter, or room temperature. Therefore —



$$V = V_1 + V_2$$

$$= 0.88 \text{ mV} + 2.80 \text{ mV}$$

$$V = 3.68 \text{ mV}$$

This value corresponds to T = 194°F.

To calibrate a thermocouple for use at reference temperatures other than 32°F often is desirable. Example D illustrates this procedure.

EXAMPLE D: CALCULATION OF THERMOELECTRIC VOLTAGE WITH KNOWN TEMPERATURES

Given: A chromel-alumel thermocouple has $T_R = 212°F$ and $T = 150°F$.

Find: The thermoelectric voltage V_2 generated.

Solution: From Table 3, V = 4.10 mV between 32°F and 212°F, and V = 2.66 mV between 32°F and 150°F. (Refer back to sketch in Example B.) Therefore —

$$V = V_1 + V_2$$

$$V_2 = V - V_1$$

$$= 2.66 \text{ mV} - 4.10 \text{ mV}$$

$$V_2 = -1.44 \text{ mV}$$

The sign reversal indicates that the test junction at 150°F is cooler than the reference junction at 212°F.

Figure 2 shows two separate thermocouples with test and reference junctions at the same temperatures T and T_R. One thermocouple is composed of metals A and C; the other, of metals B and C. For a thermocouple composed of metals A and B, thermoelectric emf is developed per Equation 1.

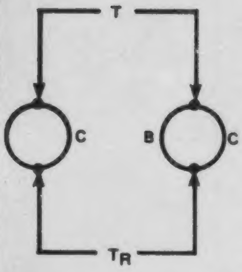


FIG. 2

$$V_{AB} = V_{AC} - V_{BC} \qquad (1)$$

where

V_{AB} = thermoelectric emf developed in a thermocouple of materials A and B.

V_{AC} = thermoelectric emf developed in a thermocouple of materials A and C.

V_{BC} = thermoelectric emf developed in a thermocouple of materials B and C.

This result enables the tabulation of V as a function of T for any thermocouple made of two standard thermocouple metals. Example E illustrates the calibration of such a nonstandard thermocouple.

EXAMPLE E: CALIBRATION OF NONSTANDARD THERMOCOUPLE

Given: A copper-constantan thermocouple with $T_R = 32°F$ and $V = 1.75$ mV at $T = 110°F$, and an iron-constantan thermocouple with $T_R = 32°F$ and $V = 2.23$ mV at $T = 110°F$.

Find: The thermoelectric emf generated across an iron-copper thermocouple at $T = 110°F$ referenced at 32°F.

Solution: A = iron (Fe), B = copper (Cu), C = constantan.

$$V_{Fe-Cu} = (V_{Fe-C}) - (V_{Cu-C})$$

$$= 2.23 \text{ mV} - 1.75 \text{ mV}$$

$$V_{Fe-Cu} = 0.48 \text{ mV}$$

Even when thermocouple tables are unavailable, there is a way to compute the thermal emf of a thermocouple, provided the temperatures to be measured are not higher than a few hundred degrees. The thermoelectric emf for a thermocouple consisting of lead (Pb) and any metal M, referenced to 0°C can be calculated from Equation 2.

$$V_{M\text{-}Pb} = aT + \tfrac{1}{2}bT^2 \qquad (2)$$

where

$V_{M\text{-}Pb}$ = thermoelectric voltage.
a and b = constants.
T = temperature of the test junction in degrees Celsius.

Table 4 lists values for the constants a and b for several metals. Example F illustrates how the thermoelectric emf for an arbitrary thermocouple is found by using Equations 1 and 2.

TABLE 4

Metal M	a (μV/°C)	b (μV/°C²)
Aluminum (Al)	-0.47	0.003
Bismuth (Bi)	-43.7	-0.47
Copper (Cu)	2.76	0.012
Gold (Au)	2.90	0.0093
Iron (Fe)	16.6	-0.030
Nickel (Ni)	19.1	-0.030
Platinum (Pt)	-1.79	-0.035
Silver (Ag)	2.50	0.012
Steel	10.8	-0.016

EXAMPLE F: CALCULATION OF THERMOELECTRIC VOLTAGE

Given: An iron-gold thermocouple has T_R = 0°C (32°F) and T = 100°C (212°F).

Find: The thermoelectric emf V.

Solution: For iron (Fe), a = 16.6 μV/°C and b = -0.030 μV/°C². For an iron-lead thermocouple —

$$\begin{aligned} V_{Fe\text{-}Pb} &= aT + \tfrac{1}{2}bT^2 \\ &= (16.6\ \mu V/^\circ C)(100^\circ C) + \tfrac{1}{2}(-0.030\ \mu V/^\circ C^2)(100^\circ C)^2 \\ &= 1510\ \mu V \\ V_{Fe\text{-}Pb} &= 1.510\ mV \end{aligned}$$

For gold (Au), $a = 2.90\ \mu V/^\circ C$, and $b = 0.0093\ \mu V/^\circ C^2$.

$$
\begin{aligned}
V_{Au-Pb} &= aT + \frac{1}{2}bT^2 \\
&= (2.90\ \mu V/^\circ C)(100^\circ C) + \frac{1}{2}\ (0.0093\ \mu V/^\circ C^2)(100^\circ C)^2 \\
&= 336.5\ \mu V \\
V_{Au-Pb} &= 0.3365\ mV
\end{aligned}
$$

From Equation 1,

$$
\begin{aligned}
V_{Fe-Au} &= (V_{Fe-Pb}) - (V_{Au-Pb}) \\
&= 1.510\ mV - 0.3365\ mV \\
V_{Fe-Au} &= 1.1735\ mV
\end{aligned}
$$

A thermopile consists of many thermocouples connected in series. The thermoelectric voltage of the thermopile is the sum of the individual emfs and, therefore, is a larger voltage, much easier to measure. Thermopiles can be used for detecting and measuring low-level radiant energy, such as that from stars. Example G demonstrates how the temperature is determined from thermopile voltage.

EXAMPLE G: THERMOPILE VOLTAGE

Given: The thermoelectric voltage across a thermopile composed of five chromel-constantan thermocouples in series is 0.1 V with reference to 0°C.

Find: The temperature of the test junction.

Solution: The thermoelectric voltage of a single thermocouple is as follows:

$$
\begin{aligned}
V &= 0.1\ V \div 5 \\
&= 0.02\ V \\
V &= 20.0\ mV
\end{aligned}
$$

The temperature corresponding to 20.0 mV in appropriate calibration tables for chromel-constantan is 287°C.

When thermocouple wires are chosen, the calibration tables should be inspected to determine which type covers the range of temperatures to be measured and which is most sensitive to changes in temperature in the range; that is, which exhibits the largest change in thermoelectric emf per degree change in temperature. The important quantity in evaluating the sensitivity of a particular thermocouple is thermoelectric power, often referred to as "thermopower." It is defined as the change in thermoelectric emf per degree change in temperature. Mathematically, this relationship is stated by Equation 3.

$$S = \frac{\Delta V}{\Delta T} \tag{3}$$

where

S = thermopower.
ΔV = change in thermoelectric voltage.
ΔT = change in temperature of the test junction.

The larger the thermopower at a given temperature, the more sensitive is the thermocouple. Table 5 compares thermopowers of several types of thermocouples at 500°C (932°F) and at -100°C (-148°F). Example H demonstrates computation of thermopower.

TABLE 5. THERMOPOWERS

Type	Composition	S(mV/°C) at 500°C	S(mV/°C) at -100°C
J	Iron-Constantan	0.060	0.040
K	Chromel-Alumel	0.040	0.030
T	Copper-Constantan	*	0.030
S	Platinum vs Platinum-10% Rhodium	0.010	*

* not tabulated at this temperature

EXAMPLE H: THERMOPOWER COMPUTATION

Given: The thermoelectric voltage V for a chromel-alumel thermocouple is 20.65 mV at 500°C and is 20.69 mV at 501°C.

Find: The thermopower S.

Solution:

$$S = \frac{\Delta V}{\Delta T}$$

$$= \frac{(20.69\ \text{mV} - 20.65\ \text{mV})}{(501°\text{C} - 500°\text{C})}$$

$$S = 0.04\ \text{mV}/°\text{C}$$

SUMMARY

The typical thermocouple consists of a reference junction at a fixed, known temperature and a test junction at the temperature to be measured. The temperature is determined from published calibration tables or is calculated using one or both of the following equations:

$$V_{AB} = V_{AC} - V_{BC} \quad (1)$$

$$V_{M\text{-}Pb} = aT + \tfrac{1}{2}bT^2 \quad (2)$$

A thermopile is a series of thermocouples whose thermoelectric voltages add, producing a larger and easier-measured emf. Thermoelectric power is the change in thermoelectric emf per degree change in temperature and indicates the sensitivity of a thermocouple.

LABORATORY

THERMOELECTRIC VOLTAGE OF THERMOCOUPLES

MATERIALS

Iron, copper, and constantan thermocouple wires.
Mercury thermometer.
Millivoltmeter (potentiometer, or strip-chart recorder).
Bunsen burner.
2 beakers.
Ring stand.
Asbestos screen.
Ice.
Thermocouple tables (see references).

PROCEDURE

1. Construct a type T thermocouple from copper and constantan wires, by twisting the opposite ends of a constantan wire onto two separate lengths of copper, as shown in Figure 3. Construct a type J thermocouple from iron and constantan in the same manner. Assemble the apparatus shown in Figure 3.

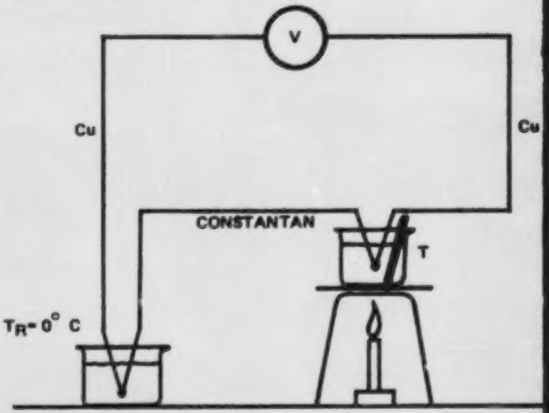


FIG. 3

2. Measure temperature of boiling water in °C with mercury thermometer, and record in Data Table. Measure thermoelectric emf of each thermocouple with millivoltmeter, and record in Data Table.

3. Determine temperature of boiling water from calibration tables for each thermocouple, and record in the results section of Data Table. If necessary, use the following equation to convert degrees Fahrenheit to Celsuis:

$$T_C = \frac{5}{9}(T_F - 32°)$$

4. Repeat Steps 2 and 3 with reference junctions and mercury thermometer at room temperature.

5. From T and J thermocouples, make an iron-copper thermocouple. Repeat Step 2, using ice water as a reference.

6. Determine thermoelectric voltage of iron-copper thermocouple from measured emf's for iron-constantan and copper-constantan (Step 2), using Equation 1:

$$V_{AB} = V_{AC} - V_{BC}$$

where

A = iron.
B = copper.
C = constantan.

Record in the results section of Data Table and compare with the measured value of Step 5.

7. Calculate thermoelectric voltage of iron-copper thermocouple, using the coefficients in Table 4 and Equations 1 and 2:

$$V_{M-Pb} = aT + \tfrac{1}{2}bT^2 \qquad (2)$$

$$V_{AB} = V_{AC} - V_{BC} \qquad (1)$$

where

A = iron.
B = copper.
C = lead.
T = temperature read from mercury thermometer in degrees Celsius.

Record under results, and compare with the values measured in Step 5 and computed in Step 6.

8. Explain any differences between the two values.

DATA TABLE

MEASUREMENTS	RESULTS
Boiling point from mercury thermometer:	
T = ________ °C	
Reference at 0°C:	
Type J (iron-constantan) V = ________ mV	T = ________ °C
Type T (copper-constantan) V = ________ mV	T = ________ °C
Reference at room temperature: T_R = ________ °C	
Type J (iron-constantan) V = ________ mV	T = ________ °C
Type T (copper-constantan) V = ________ mV	T = ________ °C
Iron-copper thermocouple V = ________ mV	V (Step 6) = ________ mV
	V (Step 7) = ________ mV

SELF CHECK

1. Sketch a typical thermocouple circuit, and briefly describe how it is used to measure temperature.

2. A thermoelectric emf of 6.92 mV is read across a copper-constantan thermocouple. Find the temperature of its test junction if its reference temperature is (a) 32°F or (b) 70°F.

3. The temperature of the test junction of a platinum-versus-platinum, 10% rhodium thermocouple is 1500°C. Find the thermoelectric voltage if the reference temperature is 0°C.

4. The test junction temperature of an iron-copper thermocouple is 300°C. For a reference temperature of 0°C, calculate the thermoelectric voltage (a) from calibration tables of iron-constantan and copper-constantan thermocouples and (b) from the equation for lead thermocouples.

5. Find the temperature of the test junctions of a thermopile consisting of five iron-constantan thermocouples with reference at 0°C, if the thermoelectric voltage is 0.1 V.

6. Find the thermopower of an iron-constantan thermocouple at T = 264°C. Use calibration tables.

10T5

THERMISTORS

The thermistor is a semiconductor transducer device whose resistance varies considerably with temperature. The resistance of some thermistors increases with a temperature increase, while that of others decreases with the same temperature increase. A thermistor exhibits characteristics that permit its straightforward application to many sensing, measurement, and control problems that otherwise would require elaborate equipment and complex circuitry. It is one of the simplest and most versatile components available for use by the electronics designer and by industry. This module explores the construction, operation, and applications of this functional component. In the laboratory, the student will measure the increase or decrease in thermistor circuits.

OBJECTIVES

Upon completion of this module, the student should be able to:

- Define the following terms:

  Thermistor.
  Self-heated mode.
  Externally-heated mode.
  Resistance-temperature characteristic.
  Voltage-current characteristic.
  Current-time characteristic.

- Describe how the thermistor is used for temperature measurement, temperature control, liquid level control, and time delay.

- Make a calibration graph for converting thermistor resistance into temperature measurements.

- Using a thermistor calibration graph, select the resistance values for five different temperatures and the temperatures for five different resistances.
- Select an application for a thermistor from several system choices.
- Select from four possibilities the correct determination of the sensitivity of a thermistor.
- Given the appropriate equipment, measure the increase or decrease in thermistor circuits, and plot resistance versus temperature for these circuits.

DISCUSSION

The thermistor [therm(o) + (res)istor] is a semiconductor device. It has the unique property that its resistance varies with temperature, increasing in some types, decreasing in others. Resistance in any electrical resistor will change as the temperature varies, but the resistance change generally is small and negligible. However, the resistance of a thermistor is unique: it changes substantially with temperature variation. Thus, the thermistor behaves as a "thermal resistor"; that is, as a resistor with a high temperature coefficient of resistance.

The temperature coefficient of resistance of a material is given by Equation 1.

$$\alpha = \left(\frac{\Delta R}{R_s}\right)\left(\frac{1}{\Delta T}\right) \qquad (1)$$

where

α = temperature coefficient of resistance.
ΔR = change in resistance when temperature is changed by an amount ΔT.
R_s = resistance of material at the reference temperature.
ΔT = change in temperature above the reference temperature.

For instance, the temperature coefficient of resistance for iron is 50×10^{-4}. For nichrome, it is 2×10^{-4}; and for silver, 38×10^{-4}. Thus, iron increases its resistance per degree rise in temperature about 25 times as much as nichrome. However, thermistors have still a much higher α value than any given here.

Notice that all the numbers given above should have units of 1/C°; that is, reciprocal Centigrade degrees. Since $\alpha = (\Delta R/R)(1/\Delta T)$, the only units that remain are associated with $1/\Delta T$, hence the units of reciprocal Centigrade degrees. A convenient way of reading the values for the temperature coefficient is simply to say "50×10^{-4} per Centigrade degree" for iron and "38×10^{-4} per Centigrade degree" for silver, and so on.

The high sensitivity to temperature causes thermistors to change their resistance by as much as 6% of their normal value for each °C temperature variation. This sensitivity makes the thermistor a useful component in temperature measurement, in temperature sensing and control, and in temperature compensation applications.

Thermistors are semiconductor devices made of ceramic materials mixed (doped) with various grades or mixtures of metalic oxides such as the oxides of manganese, nickel, cobalt, copper, iron, and uranium. After the raw materials are milled and mixed in accurate proportions, suitable binding materials are added, and the units are pressed or extruded to the desired shape.

Thermistor elements then are sintered (heated) under carefully-controlled atmospheric and temperature conditions to produce the ceramic semiconductor material. Electrode materials and lead wires then are attached. Figure 1 shows some common thermistor forms available today.

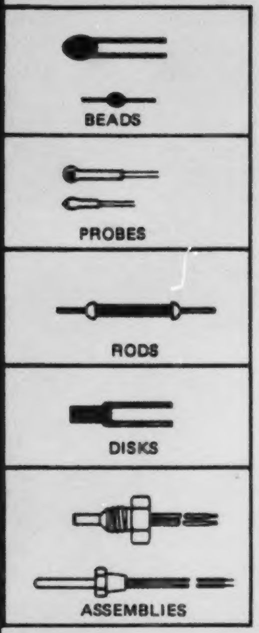


FIG. 1

THERMISTORS

Both negative-temperature-coefficient (NTC) and positive-temperature-coefficient (PTC) materials are used in the manufacture of thermistors. For negative (NTC) thermistors, the resistance decreases with temperature; for positive (PTC) thermistors, the resistance increases with temperature. Figure 2 shows the resistivity of both NTC and PTC materials, over a range of temperatures, as compared to that of a metallic conductor (platinum). Recall that resistivity is a constant for any material and is defined in Equation 2.

$$R = \frac{\rho \ell}{A} \qquad (2)$$

where

R = resistance in ohms.
ℓ = length (in convenient units).
A = cross-sectional area in the same convenient units.
ρ = resistivity, a constant characteristic of the material.

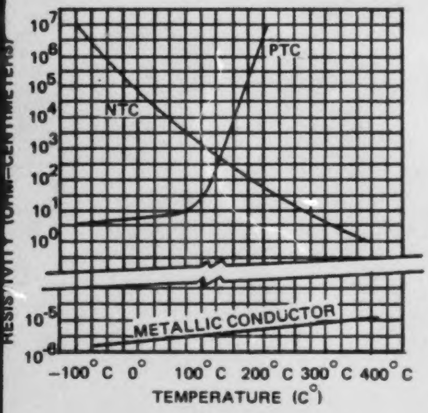


FIG. 2

Thus ρ, the resistivity, is measured in ohm-centimeters if ℓ is in centimeters and A is in $(\text{centimeters})^2$. In some types of thermistor materials, the resistance is doubled with a temperature change of as little as one degree Centigrade(°C)! PTC type thermistors have a very limited and specialized application. Since the resistance curve for PTC thermistors is abrupt and nonlinear, these thermistors can be used in motor controls and for protection against overheating.

Figure 2 also graphically illustrates the wide range of resistances available in NTC materials.

MODES OF OPERATION

There are two modes of thermistor operation — "self-heated" and "externally-heated." Externally-heated, thermistors convert changes in ambient or contact temperatures directly to corresponding changes in voltage or current flow. These thermistors are well-suited for precision temperature measurement, temperature control, and temperature compensation, because they exhibit large changes in resistance with temperature. They are widely used for applications in the range of -100°C to over 300°C.

Self-heated NTC thermistors employ the heating effect of the current flowing through them to raise and control their temperature, and, thus, to decrease their resistance dramatically. Under normal operating conditions, the temperature can rise 200 C° to 300 C°, and the resistance at high current can be reduced to 1/1000 its value at low current. This mode of operation is useful in devices such as voltage regulators, microwave power meters, gas analyzers, vacuum gages, flowmeters, and automatic volume and power-level controls. By careful circuit design, almost any desired time lag can be introduced in the response of self-heated thermistors, thus making them useful in time-delay and surge-suppression applications.

CHARACTERISTICS OF THERMISTORS

Three important characteristics of thermistors make them useful in electronic and electrical circuits:

(1) Resistance-temperature characteristic.
(2) Voltage-current characteristic.
(3) Current-time characteristic.

RESISTANCE-TEMPERATURE CHARACTERISTIC

Since thermistors exhibit a large change in resistance when subjected to a small increase or decrease in temperature, they are very accurate temperature-sensing devices. Thermistors can be placed in remote locations, because long electrical-wire leads can be used on the ohmmeter. For convenience, the ohmmeter scale usually is calibrated to read temperature directly.

VOLTAGE-CURRENT CHARACTERISTIC

Self-heating thermistors can be used to control current in a circuit. In these thermistors, a heavy current flow causes an increase in resistance as the thermistor heats up. This increased resistance lowers the current and tends to keep it constant. The current can be set at any desired level by using a variable resistor that can be calibrated to the current desired.

CURRENT-TIME CHARACTERISTIC

Since time is required for thermistors to heat up, they can be used for time-delay switches. Time delay can be controlled by means of a variable resistor calibrated in seconds (or other time units).

THERMISTOR APPLICATIONS

As there are so many circuits in which thermistors are used to great advantage, only a few of the more common applications are described.

TEMPERATURE-MEASUREMENT

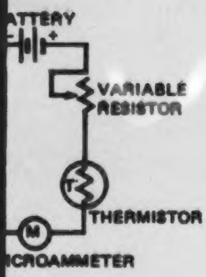


FIG. 3

The simple circuit for temperature measurement illustrated in Figure 3 consists of a battery, a variable resistor, a thermistor, and a microammeter. As the temperature changes, the resistance of the thermistor changes; and the current flow through the meter can be calibrated in terms of temperature. The resistor can be adjusted to compensate for battery age. In this circuit, the thermistor can be mounted a great distance from the meter, and ordinary copper wire can be used for the connections.

TEMPERATURE CONTROL

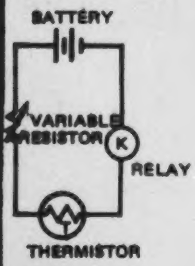


FIG. 4

Simple temperature control can be made by placing a thermistor in series with a relay, a battery, and a variable resistor, as illustrated in Figure 4. When the current is large enough, the relay will open (or close). The current can be adjusted by the variable resistor to actuate the relay at the desired thermistor temperature.

LIQUID LEVEL CONTROL

If a thermistor is placed in series with a relay and a battery, the relay can be made to operate if the thermistor is suspended in air (higher temperature) and to drop out when the thermistor is immersed in liquid (lower temperature). This situation is shown in Figure 5. In air, the NTC thermistor heats up, causing the resistance to drop and permitting enough current to flow to pull in the relay. If the thermistor is submerged in a liquid, it will be cooled, causing its resistance to increase. As a result, the current through the relay is reduced to a point at which the relay is no longer actuated. This device can be used as a liquid-level indicator or as a liquid-level control device.

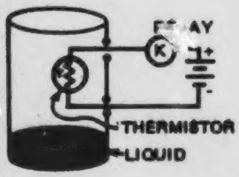


FIG. 5

TIME DELAY

If a thermistor is placed in series with a variable resistor, a battery, and a relay as shown in Figure 6, a variable time-delay relay can be made. When the switch is closed, current flow is limited by the high resistance of the thermistor, until it heats up and permits sufficient current to flow to close the relay. An increase in circuit resistance increases the time delay. With proper selection of components, time delays from milliseconds to many seconds can be produced.

TESTING OF THERMISTORS

The equipment used to test thermistors must possess a measurement capability corresponding to the precision desired in measuring thermistor characteristics. Where applications or trouble-shooting measurements are made, any instrument that has a low-power ohm-function can be used. A digital VOM or a FET VOM is suggested for use, rather than the common Triplett or Simpson meters which contain a 1.5-volt battery. Current from a 1.5-volt battery can cause self-heating of the thermistor, and this will make the readings inaccurate.

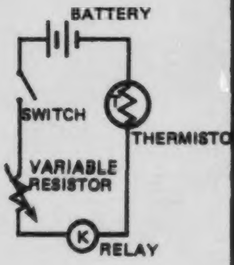


FIG. 6

SUMMARY

A thermistor is a resistive element whose resistance changes substantially in response to its temperature. It is a semiconductor device constructed of ceramic materials doped with various mixtures of metallic oxides. Some thermistors are self-heated, or heated by the current through them, and often are used to control current flow or to produce a time delay. Others are heated externally, or heated by ambient or contact heat, and are employed in temperature-measurement and control systems. Though simple, thermistors perform valuable functions, which if duplicated by other components would require elaborate equipment and complex circuitry.

RESISTANCE/TEMPERATURE CHARACTERISTICS OF A THERMISTOR

MATERIALS

VOM.
Thermistor.
Ice.
Boiling water.
Two beakers.
Bunsen burner, ring stand, and wire gage.
Supply of hot, cold, and mixed tap water.
Laboratory thermometer.

PROCEDURE

In this experiment, the student will make a resistance-temperature graph for a thermistor in order that it can be used to measure temperature.

1. Connect VOM leads to thermistor, and record resistance of thermistor at room temperature (in air). In Data Table 1, record room temperature (as shown on a laboratory thermometer).
2. Place thermistor and thermometer in a container of melting ice. Record resistance and temperature.
3. Place thermistor and thermometer in a container of boiling water. Record resistance and temperature.
4. Record resistance and temperature for thermistor and thermometer in tap water from cold water tap, from hot water tap, and from both taps mixed.
5. Record type of thermistor and type of VOM.
6. Plot resistance versus temperature on the graph in Data Table 2. Carefully draw a smooth curve, connecting the data points plotted and extending the curve to the sides of the graph. (Semi-log axis is used for resistance values, because resistance covers such a wide range of values.)
7. Explain the significance of the data plotted and some industrial applications of this thermistor.

RESISTANCE CHANGES OF A THERMISTOR

MATERIALS

Thermistor (same as in first experiment).
Ice.
Boiling water.
No. 1449 bulb and socket.
d.c. power supply or battery (variable-voltage capacity).
Two beakers.
Bunsen burner, ring stand, and wire gage.
Thermometer.

PROCEDURE

Assemble circuit shown in Figure 7, using a No. 1449 bulb and socket, a thermistor, and a variable-voltage power supply.

Note brilliance of the lamp at room temperature. Expose thermistor to cold (touch it to the ice water) and then to hot (touch it to the boiling water). At each temperature extreme, note the lamp brilliance. From these observations, analyze the operation of the thermistor in this circuit. CAUTION: Use a voltage as low as possible to light the bulb with the thermistor at room temperature. Do not change this voltage setting during the experiment.

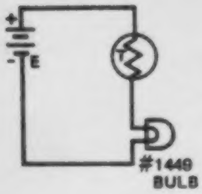


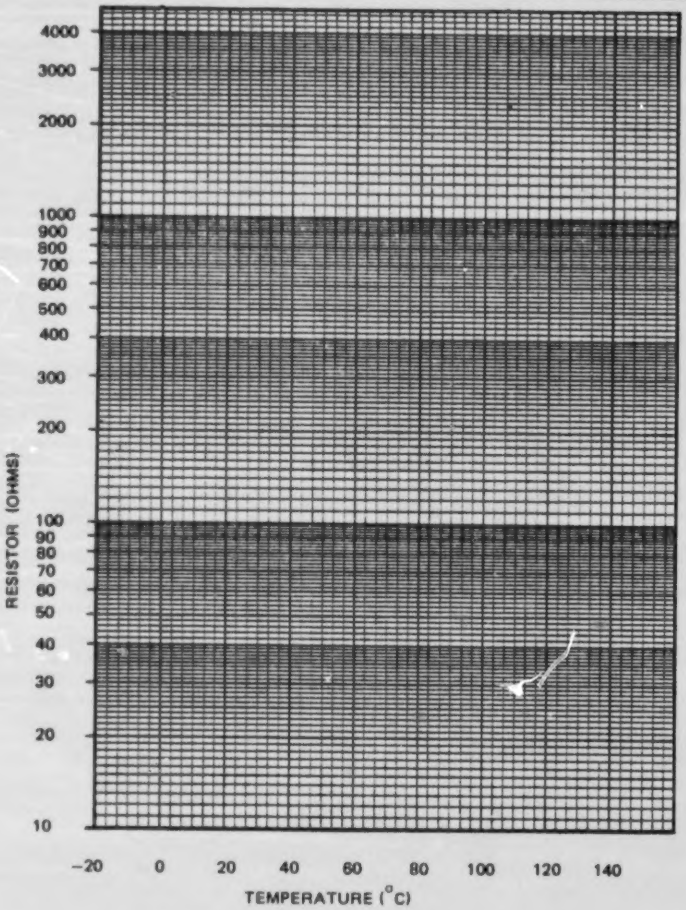
FIG. 7

DATA TABLE 1

Situation	Thermometer Reading (°C)	R (Ohms)
Room temperature		
Ice water		
Boiling water.		
Cold tap water		
Hot tap water		
Mixed tap water		

Type of thermistor ______________________________.
Type of VOM ____________________________________.

DATA TABLE 2



TYPE OF THERMISTOR________ TYPE OF VOM ________

DATA TABLE 3

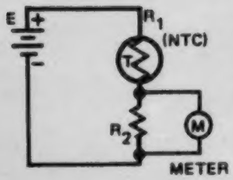
Temperature	Resistance (Ω)	Resistance	Temperature (°C)
15°C		500 Ω	
80°C		1500 Ω	
40°C		900 Ω	
25°C		2000 Ω	
60°C		1250 Ω	

SELF CHECK

1. Define these terms:

   Thermistor.
   Self-heated mode.
   Externally-heated mode.
   Resistance-temperature characteristic.
   Voltage-current characteristic.
   Current-time characteristic.

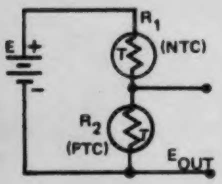
2. Describe how the thermistor is used for temperature measurement, temperature control, liquid level control, and time delay.



3. Using the thermistor calibration graph drawn in the laboratory, fill in Data Table 3.

4. In the adjoining circuit, how will the voltage across resistor R_2 vary if the temperature decreases?

   Increases?
   Decreases?
   Remains the same?



5. In the adjoining circuit R_1 and R_2 are thermistors with resistances at room temperature and identical characteristics. The only difference is that one has an NTC, and the other has a PTC. If the temperature increases, how will the output voltage E_{out} change?

   Increases?
   Decreases?
   Remains the same?

6. In a positive temperature thermistor, a temperature increase causes what change in the resistance?

   Increases?
   Decreases?
   Remains the same?

7. Which of the following would not be a possible application for a thermistor?

   Electronic thermometer.
   Fire detection and warning system.
   Fuse.
   Temperature control for PTC components.

8. The sensitivity of a thermistor is determined by which of the following?

Its rate of power dissipation.
Its reaction to self-heating effects.
Its ability to respond to a wide temperature range.
The amount of resistance change for a given temperature variation.

9. Find the temperature coefficient of resistance of a metal wire if the resistance of the strip at room temperature (295K) is 1Ω and at 370K is 1.001Ω.

10. What length of a wire will have a resistance of 10Ω if the cross-sectional area of the wire is 0.1 cm^2 and the resistivity of the wire is 0.1 Ω-cm.

11. Find the resistivity of a wire that is 10 m long, has a cross-sectional area of 1×10^{-4} m^2 and a resistance of 10Ω.

10T6

PYROMETERS

A body when heated emits thermal radiation in the infrared or visible parts of the spectrum. Measurement of either total power radiated or color associated with thermal radiation provides a noncontact method for determining temperature of the hot body. A device known as a pyrometer is especially useful in making appropriate measurements if the body temperature is high.

In this module, the student will use pyrometers to perform noncontact temperature measurements.

OBJECTIVES

To complete this module, the student should be able to:

- Define emissivity, Stefan-Boltzmann constant and blackbody radiation.
- State mathematical relationship between temperature of a heated body and emission of thermal radiation and explain each term, as well as indicate how this relationship can form the basis for a radiation pyrometer.
- Describe relationship between color of a heated body and its temperature and indicate how this relationship can form the basis for an optical pyrometer.
- Describe operation of a radiation pyrometer and an optical pyrometer.
- Use a pyrometer to measure temperatures of heated bodies.

DISCUSSION

The word pyrometer can be applied to many different types of temperature-measuring devices, including thermocouples. In this module, however, we shall adopt the more restricted definition of pyrometer. We shall consider pyrometers to be noncontact temperature-measuring devices that respond to certain characteristics of electromagnetic radiation emitted by the object whose temperature is to be measured. Further, we shall use pyrometers as devices for measurement of relatively high temperatures. For measurement of lower temperatures, other devices such as thermocouples or resistance thermometers would be more appropriate. Clearly, when temperature becomes high, noncontact or pyrometric methods are more appropriate.

The physical basis for the operation of pyrometers is found in emission of radiation by all heated bodies. Any heated body emits thermal radiation. Such emission is a form of electromagnetic radiation, similar in nature to radio waves or light waves. Figure 1 shows thermal radiation emitted by a certain body that has been heated by passing electric current through a resistor embedded in the body.

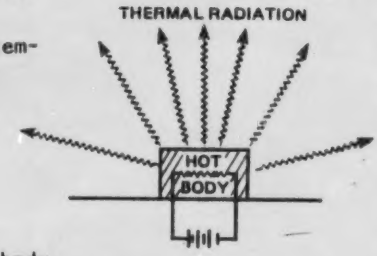


FIG. 1

The amount of radiation emitted increases with body temperature according to the following equation:

$$P = \sigma A \varepsilon T^4$$

where

P = energy per unit time or power (watts) emitted by body.
T = absolute temperature, degrees Kelvin.
A = radiating surface area of heated body, cm^2.
σ = constant called the Stefan-Boltzmann constant, which is numerically equal to 5.669×10^{-12} $w/cm^2\text{-deg}^4$.
ε = surface parameter called emissivity.

Emissivity (a dimensionless number) is a quantity that describes emissive power of the surface. It is the ratio between radiant energy emitted by the surface of the body in question and radiant energy that would be emitted by the surface of a perfect radiator at the same temperature.

Emissivity can have any value between 0 and 1. In the limiting case of $\varepsilon = 1$, we have what is called a blackbody. A blackbody at any specified temperature emits the maximum possible radiant power from a given area. It can thus be considered a perfece emitter of radiant energy. In actual practice, there is no real substance that behaves like a perfect blackbody. Some substances, however, such as soot or lampblack, are reasonably close approximations, since they have values of emissivity close to 1. Many other practical substances have values of emissivity that are considerably lower.

Table 1 lists values of emissivity for some substances of interest. Shiny, unoxidized metallic surfaces typically have quite low values of emissivity, much smaller than 1. Normally, when the surface of a metal is oxidized, emissivity increases. Emissivity may also depend on temperature. Note that some materials listed have values of emissivity that are not much less than 1. Examples A, B and C show how Equation 1 applies to different radiant emission calculations.

TABLE 1. EMISSIVITY FOR SEVERAL MATERIALS

Material	(Dimensionless) Emissivity
Aluminum (unoxidized, 300°K)	0.022
Nickel (unoxidized, 300°K)	0.045
Tungsten (unoxidized, 300°K)	0.024
Lead (unoxidized, 300°K)	0.05
Tin (unoxidized, 300°K)	0.043
Steel (unoxidized, 300°K)	0.08
Lead (500°K)	0.08
Tin (500°K)	0.11
Steel (oxidized)	0.80
Granite	0.44
Hard rubber	0.92
Porcelain	0.92

Radiation emitted by heated bodies is spread over a broad range of wavelengths, covering microwave, infrared and, in some cases, visible regions of the spectrum. For comparison purposes, the visible portion of the electromagnetic spectrum lies at wavelengths between 0.4 and 0.7 micrometers. At wavelengths greater than 0.7 micrometers, there is infrared radiation, which is invisible to the human eye but which may be detected by the sensation of warmth that it produces on the human skin.

Details of wavelength distribution of thermal radiation are beyond the scope of this module. We simply illustrate (Figure 2) the manner in which the spectral distribution of radiant power changes with increasing temperature, over the spectral range cov-

EXAMPLE A: RADIANT POWER CALCULATION

Given: Equation 1 and an object with $\varepsilon = 1$ and with surface area 10 cm^2 at a temperature of 727°C.

Find: Amount of radiant power emitted.

Solution: Absolute temperature is 727 + 273, or 1000°K. From Equation 1,

$$P = \sigma A \varepsilon T^4, \text{ or}$$

$$\text{Power} = 5.669 \times 10^{-12}\ w/cm^2\text{-deg}^4 \times 10\ cm^2 \times 1 \times (1000\ \text{deg})^4$$

$$P = 56.69 \text{ watts.}$$

EXAMPLE B: EMISSIVITY CALCULATION

Given: A body at temperature 500°K emits radiant power density of 1.17 $watts/cm^2$.

Find: Emissivity.

Solution:

Since $P = \sigma A \varepsilon T^4$, we have

$$P/A = 1.17\ w/cm^2 = \varepsilon \times 5.669 \times 10^{-12}\ w/cm^2\text{-deg}^4 \times (773\ \text{deg})^4$$

$$1.17 = 2.02\ \varepsilon.$$

Thus:

$$\varepsilon = 1.17/2.02 = 0.58.$$

EXAMPLE C: RADIATION EMISSION CALCULATION

Given: Absolute temperature of a body increases 10%.

Find: Percentage of radiant emission increase.

Solution: Radiant emission increases by a factor of $(1 + 0.1)^4 = 1.464$, an increase of 46.4%.

ering visible and near infrared portions of the spectrum. Spectral radiant intensity is plotted in relative units and is power per unit surface area per unit wavelength interval, as a function of wavelength. Note that total radiant power (the area under the curve) increases very rapidly as temperature increases.

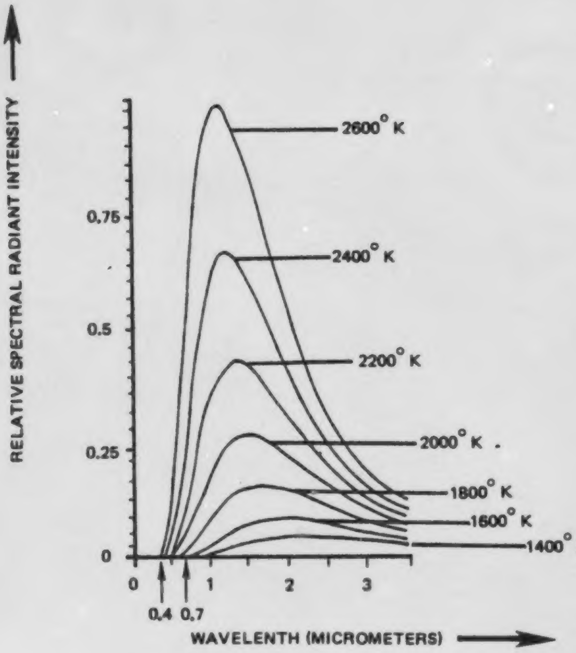


FIG. 2

This is in agreement with the fourth power law expressed by Equation 1. In addition, spectral distribution changes with the peak wavelength shifting toward shorter wavelengths as temperature increases. Thus, for a temperature of 1400°K, there is very little radiant energy within the visible region (0.4 to 0.7 micrometers), whereas, for a temperature of 2600°K, a considerable amount of energy falls within this region.

The color of a heated body as viewed in the emitted radiation is dependent on body temperature. For a heated body to be visible because of its emitted radiation, some radiation with wavelength in the visible region must be present. As temperature is increased gradually, the body first glows with a dull red color, then a bright red and eventually a bluish white color. Color changes because the radiation includes relatively more energy at shorter wavelengths as temperature increases. Table 2 illustrates progression of color change as temperature increases.

TABLE 2. APPARENT COLOR OF HEATED BODIES

Temperature	Color
475°C	Dull red
700°C	Dark red
900°C	Bright red
1100°C	Yellowish-red
1200°C	Yellow-white
1400°C	White
1550°C	Blue-white

In measurement of radiation from heated bodies, one is interested in limited intervals of wavelength and may consider emissivity as a function of wavelength because, for many practical cases, emissivity does change considerably over the total wavelength interval within which the body is emitting. Thus, one has the concept of an emissivity for unit wavelength interval, that is, a spectral emissivity, in contrast to the total emissivity defined earlier. Spectral emissivity is the ratio of power per unit area emitted by the body within a narrow spectral interval divided by power per unit area emitted within the same narrow spectral interval by a blackbody at the same temperature. Table 3 presents some values of spectral emissivity at the wavelength of 0.65 micrometers, a wavelength often used for measurement.

TABLE 3. SPECTRAL EMISSIVITY

Material	Spectral Emissivity Values for Wavelength 0.65 μm
Carbon (unoxidized)	0.85
Nickel (unoxidized)	0.36
Tungsten (unoxidized)	0.43
Steel (unoxidized)	0.35
Steel (oxidized)	0.80
Aluminum oxide	0.30
Black felt	0.86
Tungsten (1400°K)	0.453

Thermal radiation emitted by heated bodies leads to a method for determining temperature of the body through a measurement of the amount of radiation. This measurement does not require contact between the measuring device and the hot body so that high temperatures, or temperatures of moving or inaccessible objects may be determined. There are basically two types of measurements that may be performed, taking advantage either of total amount of radiation emitted, according to Equation 1, or of dependence of spectral characteristics of the radiation on the hot body temperature.

The first type of measurement, which essentially measures total radiant emission, uses a device called a radiation pyrometer. Basic structure of a common type of radiation pyrometer is shown in Figure 3. Since one knows the relation between total radiation from the surface of the body and its temperature, measurement of the radiation is enough to determine temperature. In Figure 3, we see radiation collected by a concave mirror and focused on a radiation-sensing element, indicated in this figure

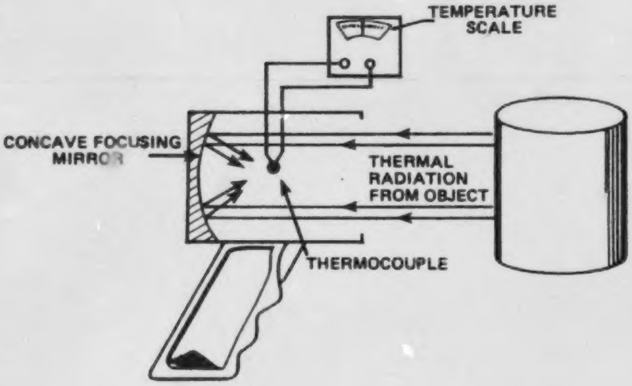


FIG. 3

as a thermocouple. The radiation is collected and heats the thermocouple. Heating of the thermocouple by the radiation leads to generation of a voltage as described in Module 10-0. Therefore, temperature increase of the thermocouple is proportional to radiation received, which, according to Equation 1, is proportional to the fourth power of the temperature of the heated body. One must also include a subtractive term proportional to the fourth power of the surrounding background temperature, typically around 300°K, because the thermocouple also absorbs background radiation. In many cases of high temperature, this term will be negligible.

Over a reasonable range, measurement does not depend on distance between pyrometer and hot body. As one moves farther from the body, the amount of radiation collected by the pyrometer decreases. Size of the image of the hot body formed by the mirror also decreases in the same ratio. Since the image is always larger than the sensor, power received by the sensor will not change and measured value will be independent of distance.

Practical use of a radiation pyrometer depends on knowing the emissivity of the surface. If surface emissivity is less than 1, apparent temperature will be lower than actual temperature because the pyrometer receives less radiation than would be emitted by a blackbody. Figure 4 shows relation between true temperature and observed temperature for various values of the emissivity. For emissivity equal to 1, the observed temperature equals the true temperature (see dashed lines in Figure 4, for example, showing both true and observed temperature equal to 1000°C). For values of emissivity less than 1, 0.6, for example, true temperature (2000°C, for example) is higher than observed temperature (1700°C, for example), with difference becoming greater as emissivity becomes smaller. Thus, for practical use of the radiation pyrometer, one must have knowledge of the emissivity of the surface.

Requiring knowledge of the emissivity can be a source of uncertainty in the measurement because emissivity depends on many conditions, including temperature, surface roughness, particle size for ceramic materials and state of surface oxidation. Moreover, surface oxidation can vary with time for materials heated to high temperatures. These factors mean that temperature measurement with radiation pyrometers can have a fairly large source of error, unless careful calibration against surfaces of similar structure is carried out at known temperatures (Example D).

An optical pyrometer measures thermal energy emitted within a narrow wavelength interval by matching apparent color of the object to color of a filament heated by an electrical current. This measurement makes use of the fact that apparent color of a heated body changes with temperature and that, within a given narrow wavelength interval, the amount of radiation emitted increases with temperature. The optical pyrometer is illustrated in Figure 5. It is similar to a low-power telescope with the filament located in the common focal plane of the two lenses. A piece of red glass is located between lamp and eyepiece. The red glass transmits light only in a narrow spectral interval, usually chosen to be around 0.65 micrometers. When the telescope is focused on a heated glowing object, the image of the target is formed. The observer, view-

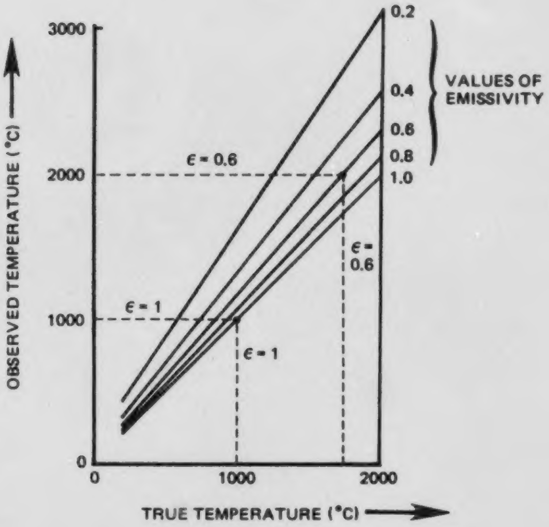


FIG. 4

EXAMPLE D: TRUE TEMPERATURE CALCULATION

Given: A material with ε = 0.5 and temperature observed with a radiation pyrometer of 600°C.

Find: True temperature.

Solution: Using Figure 4, and interpolating halfway between the curves for ε = 0.4 and ε = 0.6 at 600°C, we obtain a true temperature of 760°C.

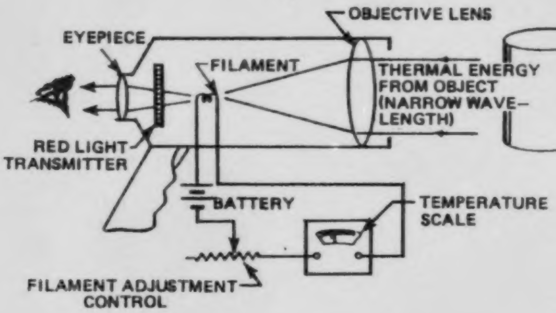


FIG. 5

ing through eyepiece and red glass, sees the image of the lamp filament superimposed on the image of the heated body. By adjusting current to the lamp filament, brightness of the lamp filament is varied until it matches that of the target and seems to disappear against the background of the target. Thus, this type of operation is sometimes called a disappearing filament optical pyrometer.

Temperature is determined by a measurement of current. The pyrometer can be calibrated by viewing a blackbody target at known temperatures. Thus, the meter reading of the curreht in the matched condition can be calibrated to read a blackbody temperature directly.

As with the radiation pyrometer, the effect of varying emissivity is important. The optical pyrometer gives an observed value of temperature that is equal to the true value of temprature only if the spectral emissivity is equal to unity. If spectral emissivity in the narrow wavelength range in which the measurement is carried out is less than unity, observed temperature will be lower than true temperature and a correction must be applied. The correction is made with a table or graph in the same manner as was described for the radiation pyrometer. A typical set of correction curves is presented (Figure 6).

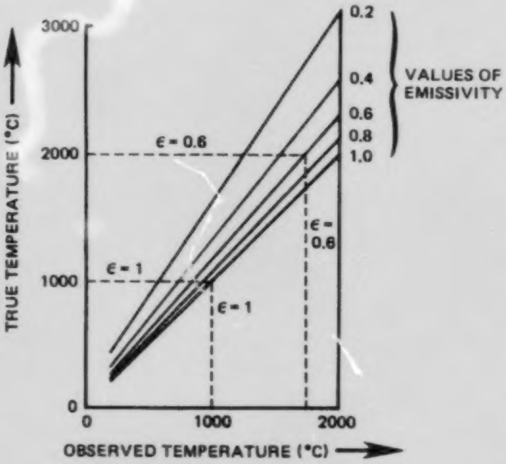


FIG. 6

Because the optical pyrometer requires the body to be glowing visibly, it is useful at a higher range of temperature than is the radiation pyrometer. The low temperature limit for practical measurements with an optical pyrometer is around 700 to 800°C. Above target temperatures of 1300 or 1400°C, an additional absorbing glass filter can be placed between the filament and the objective lens. This reduces apparent brightness of the target to a level that can again be matched by the filament and read on the meter recalibrated for the higher temperature range. Thus, the range of the optical pyrometer can be extended to higher temperature values.

LABORATORY

MATERIALS

Radiation pyrometer (model HSA-4E from William Wahl Company, or equivalent).
Bunsen burner.
Crucible.
Stand and brackets to support crucible.
Lead and tin metal.
Optical pyrometer (Leeds and Northrup series 8627 or equivalent).
Tube furnace capable of reaching 1000°C or more (Lindberg model 55035 or equivalent); furnace should have temperature sensor and display).
Tungsten metal sheet.

PROCEDURE

The procedure presented will assume use of pyrometers as described in the materials list. If different types of pyrometers are employed, instructions may vary somewhat. Consult instructor and instructions supplied by instrument manufacturer and follow their guide.

1. Set crucible on stand; add metallic lead. Heat lead with bunsen burner until lead begins to melt. When material is partially melted, and liquid and solid lead are present simultaneously, material should be at a temperature of melting point of lead.
2. Sight radiation pyrometer on sample. Make sure that sample fills field of view of pyrometer. Set pyrometer dial to known emissivity of lead. (Use Table 1 for emissivity at 500°K.) Obtain temperature from meter reading. (For models of pyrometer for which emissivity cannot be dialed in, use Figure 4 to correct observed temperature and obtain true temperature.)
3. Repeat above procedures with tin metal in crucible.
4. Compare temperatures obtained to known melting points of lead and tin, 327 and 232°C, respectively.
5. Set tungsten metal sheet within tube furnace and turn furnace on. Monitor temperature with thermocouple sensor built into furnace.
6. For some value of temperature in furnace at which sample has begun to glow, sight optical pyrometer on sample. Vary filament current until filament disappears against background of sample. Observe temperature reading on pyrometer display.

7. Increase temperature in furnace; repeat step 6. Obtain temperature measurements with pyrometer at several different furnace temperatures.
8. Correct temperature readings obtained with pyrometer, using spectral emissivity of tungsten from Table 3 for 1400°K spectral emissivity. (Use manufacturer's calibration curves. If these are not available, use Figure 6.)
9. Compare temperature readings obtained with pyrometer with those indicated on the furnace's meter. Discuss possible sources of difference.

VIBRATIONS AND WAVES

APPLICATION MODULES

		Page
11M3	Vibration Isolation	353
11M4*	Resonance	363
11F1*	Cavity Vibrations and Standing Waves	374
11F2	Doppler Effect	382
11E1*	Alternating Voltage and Current	398
11E2*	Phase Relationships in AC Circuits	409
11E3	Oscilloscope	425
11E4*	Rectifiers	437

* Identifies concept support modules that introduce new physics.

VIBRATION ISOLATION

Concept Module 11-0, "Vibrations and Waves," explains how energy is transferred from one place to another via waves. In some systems, every attempt is made to maximize energy transfer via waves. For example, stereo sets and musical instruments are designed to provide maximum transfer of wave energy to the listener. On the other hand, mechanical waves or vibrations may be undesirable in some instances, and attempts are made to minimize their effects.

Mechanical vibrations are produced by almost every kind of machinery. Electric motors, internal combustion engines, air compressors, vacuum pumps, printing and manufacturing presses, and fans all produce unwanted vibrations. If these vibrations are not eliminated or reduced to an acceptable level, they can destroy the machinery or limit its proper operation.

In this module the characteristics of steady-state vibrations will be illustrated, and the methods used to eliminate or reduce these vibrations in various kinds of machinery will be discussed. In the laboratory, the student will measure the acceleration of a vibrating system and its support to determine transmission of vibration.

OBJECTIVES

Upon completion of this module, the student should be able to:

- Distinguish between vibrations produced by internal and external sources. Give at least two examples of each type, including a labeled drawing of each.
- Define the following terms:

Vibration.
Periodic.
Amplitude.
Transient vibration.
Shock.
Isolation.
Steady-state vibrations.
Superposition.
Vibrators.
Frequency.
Period.

- Calculate transmission of mechanical vibrations, given acceleration values of machinery and support structure.

- Given the appropriate equipment, measure the accelerations of a vibrating system and its support, and determine the transmission of the isolating system.

DISCUSSION

DEFINITIONS

In Concept Module 11-0, "Vibrations and Waves," the word "vibration" was used to describe waves or wavelike motion. In this module, the broader meaning of vibration will be used. Vibration indicates any to-and-fro motion or oscillation of a mechanical system.

Some devices are specifically designed to cause vibration or shaking and are called vibrators. Foundries use special vibrators to loosen castings from molds. Vibrators also are used in medicine to stimulate blood circulation and muscles. All sound begins with a vibrating object; for example, when a person speaks, the vocal cords initiate the vibration that we recognize as speech sounds.

On the other hand, vibrations that weaken structures and produce wear and tear in industrial machines certainly are undesirable. In cities, trucks, subways, elevated trains, and other traffic cause considerable vibrations that are transmitted effectively throughout the steel structure of a large building. Engineers deal with this problem by using special vibration-absorbing materials. Pneumatic tires are used on automobiles in an effect to reduce undesirable vibrations created by running over rough surfaces. Automobile engines are mounted on rubber supports to reduce vibration; and more cylinders have been added to engines to smooth out engine vibration, resulting in a smoother flow of power.

Many industries deal with the problem of vibration and its effect on machinery. High-speed, rotating parts produce vibration unless they are accurately balanced. The supports of such rotating parts were once bolted securely to the machine frame or its foundation. Today the supports for high-speed, rotating parts are mounted on flexible connections, such as springs or rubber, thereby reducing the effect of vibration in machines.

Mechanical vibrations have various frequencies and amplitudes. Frequency is the number of complete vibrations made during a certain time period. Amplitude is the maximum distance the item travels from its normal position of equilibrium or rest. The period is the time required to make one complete vibration or oscillation, and is related to frequency by Equation 1.

$$T = \frac{1}{f} \qquad (1)$$

where

T = period (time) in seconds.
f = frequency in Hz.

Vibrations that repeat themselves at regular intervals are

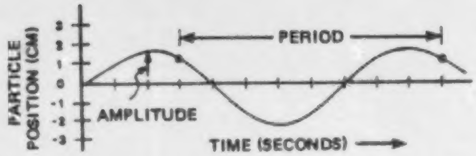


FIG. 1

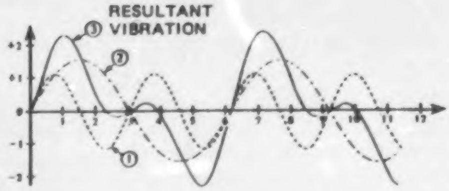


FIG. 2

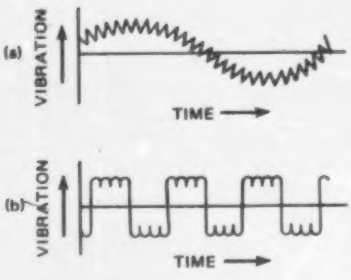
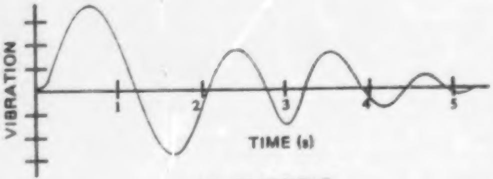
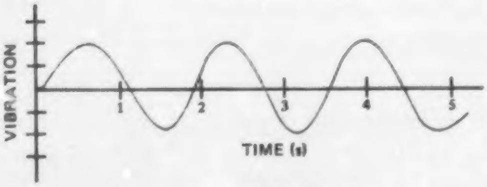


FIG. 3



(a) TRANSIENT



(b) STEADY STATE

FIG. 4

called periodic. Both amplitude and period are shown in Figure 1 for a vibration that varies as a sine wave.

Most vibrations are complex, meaning that they are the sum of two or more simple vibrations. Figure 2 illustrates the adding of vibrations 1 and 2 to obtain the resultant, or composite, vibration 3.

Figure 3 illustrates the nature of other types of complex vibrations.

If machinery vibrates in a regular manner, the vibrations are called "steady-state." These vibrations normally are found in all machines and appliances, such as washers, dryers, and electric motors. Vibrations that are sudden and die out quickly are called "transient vibrations," or "shock vibrations." Figure 4 illustrates transient and steady-state vibrations.

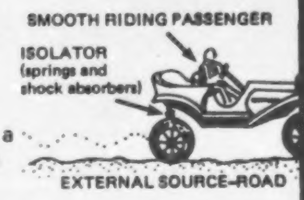

SMOOTH RIDING PASSENGER
ISOLATOR
(springs and
shock absorbers)
EXTERNAL SOURCE-ROAD

FIG. 5

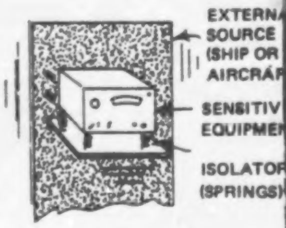

EXTERN
SOURCE
(SHIP OR
AIRCRAF
SENSITIV
EQUIPMEN
ISOLATOR
(SPRINGS)

FIG. 6

TWO CLASSES OF VIBRATIONS

EXTERNAL

Vibrations in this class are caused by sources external to the machinery or equipment. Three examples of this class are illustrated in Figures 5, 6 and 7.

Highly variable vibration of wheels running over an uneven road surface are reduced greatly by automobile springs and shock systems. The ride inside the car is relatively smooth and comfortable.

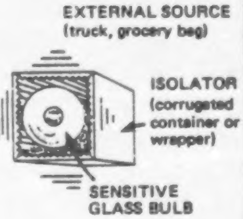

EXTERNAL SOURCE
(truck, grocery bag)
ISOLATOR
(corrugated
container or
wrapper)
SENSITIVE
GLASS BULB

FIG. 7

A delicate electronic instrument is mounted on springs to absorb vibrations produced by the motion of the bulkhead to which it is attached.

A light globe is protected from external vibrations by a corrugated container or wrapper in which it is packaged.

INTERNAL

Vibrations in this class are produced by the machine's moving parts, which cause the entire system or machine to shake. Three examples of this class of vibration are shown in Figures 8, 9 and 10.

An unbalanced automobile tire (Figure 8) can cause vibration in the entire car. This vibration can be reduced greatly when the wheel is balanced.

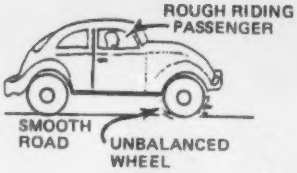

ROUGH RIDING PASSENGER
SMOOTH ROAD
UNBALANCED WHEEL

FIG. 8

Vibratory forces frequently are induced in electrical machinery by magnetic effects. These effects can be reduced if slots are placed in the armature lamination to reduce jerking caused by the rapidly-changing magnetic field (Figure 9).

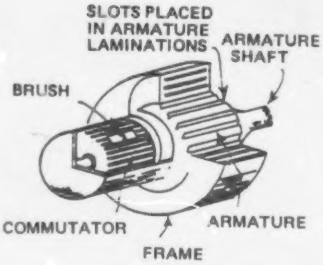

SLOTS PLACED IN ARMATURE LAMINATIONS
ARMATURE SHAFT
BRUSH
COMMUTATOR
ARMATURE
FRAME

FIG. 9

An automobile axle bearing that has become worn and does not have even tolerances (Figure 10) will cause the wheel to vibrate, particularly at high speeds. This kind of vibration usually is reduced only by replacement of the worn bearings.

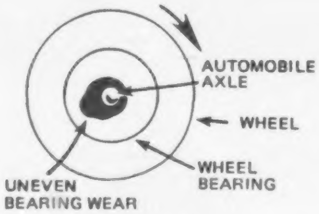

AUTOMOBILE AXLE
WHEEL
WHEEL BEARING
UNEVEN BEARING WEAR

FIG. 10

VIBRATION ISOLATION

Excessive vibration must be isolated or reduced for the protection of personnel and of the equipment itself. Effectiveness of vibration isolation provided by a system may be evaluated by measuring the amount of vibration transmitted through an isolating medium.

Transmission of a vibration-isolation system can be measured by determining accelerations on both support structures and vibrating machinery. Transmission of the system for external vibration is given by Equation 2.

$$T_E = \frac{a_m}{a_s} \qquad (2)$$

where

T_E = external transmission of a vibration isolation system.
a_m = acceleration of machinery.
a_s = acceleration of support.

For internal vibrations, transmission is defined by Equation 3.

$$T_I = \frac{a_s}{a_m} \tag{3}$$

where

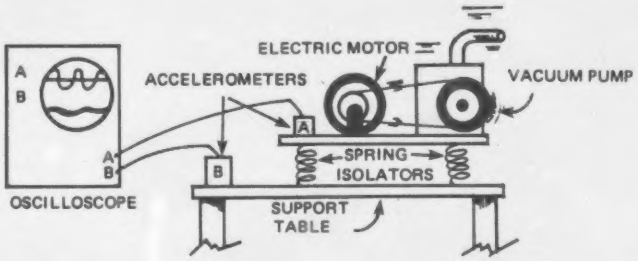
T_I = internal transmission of a vibration-isolation system.

The acceleration of the vibrating source is placed in the denominator and acceleration of the isolated structure or machinery is placed in the numerator. In either case, the smaller acceleration is always the numerator.

Example A illustrates use of Equation 3 to solve a problem

EXAMPLE A: TRANSMISSION CALCULATION

Given: A vacuum pump is mounted on springs (see sketch). Amplitude of trace A is 2 cm peak-to-peak, and peak-to-peak amplitude of trace B is 0.2 cm. (Note: The amplitude on the oscilloscope of trace A or trace B is proportional to the acceleration of the pump and table, respectively.)



Find: T_I.

Solution: Transmission of the spring isolators is —

$$T_I = \frac{a_s}{a_m} = \frac{\text{acceleration of support}}{\text{acceleration of vacuum pump}}$$

$$= \frac{0.2\ \text{cm}}{2\ \text{cm}}$$

$$T_I = 0.1$$

Springs reduce vibrations to 10% of their original value.

SUMMARY

Vibration in a mechanical system is a wavelike, to-and-fro motion. When such vibration is undesired, it can be isolated by use of mounts that absorb some, or hopefully most, of it. The difference between the internal vibration of the equipment and the external vibration of its support can be measured to calculate the effectiveness of the mounts. The transmission of vibration refers to the amount of vibration "transmitted" from source to support.

VIBRATION ISOLATION AND TRANSMISSION

MATERIALS

Accelerometer with cable. (Two accelerometers are preferable for this experiment.)

Aluminum block 1" x 2" drilled and tapped to accept accelerometer. (If two accelerometers are available, two aluminum blocks are needed.)

Oscilloscope. (If two accelerometers are available, a dual-trace oscilloscope is desirable.)

Oscilloscope camera. (This piece of equipment is not absolutely required, but would greatly simplify the experiment.)

Electric motor and adjustable counterweights as in Figure 11 (can be nuts screwed into threaded rod.)

Guard for counterweights. (See Note before first step below.)

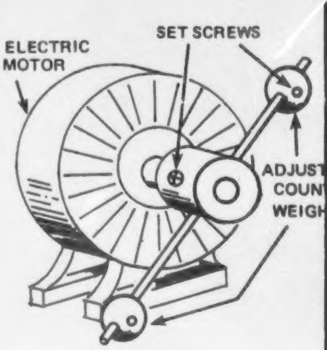


FIG. 11

PROCEDURE

Note: This experiment requires a guard to be placed over rotating weights to avoid possible accident. For simplicity no guards are shown in Figure 12. Have instructor check out system before starting experiment.

1. Set up experiment as shown in Figure 12. Have counterweights balanced and guards in place.

2. Turn on motor; measure acceleration of motor mount and support surface at points A and B. Calculate transmission: $T = \frac{a_s}{a_m}$; record in Data Table as Case 1.

3. Turn off motor; move one counterweight either toward or away from motor axis, producing a slight imbalance.

4. Repeat step 2; record results as Case 2 of Data Table. If resulting vibrations are not large enough, turn off motor and increase imbalance of counterweights.

5. Turn off motor; balance counterweights. Repeat step 2,

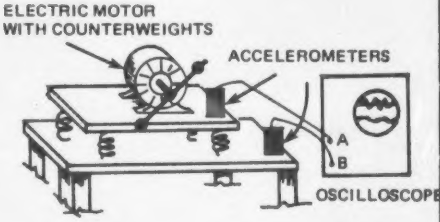


FIG. 12

recording results as Case 3 of Data Table.

6. Are vibrations produced by the motor-counterweight arrangement used in experiment section external or internal?

DATA TABLE

	a_m	a_s	$T = \frac{a_s}{a_m}$	Class of Vibration
Case 1				
Case 2				
Case 3				

SELF CHECK

1. Distinguish between vibrations produced by internal and external sources. Give two examples of each type; prepare a labeled sketch of each.

2. In one or two sentences define these terms:

   Steady-state vibration.
   Transient vibration.
   Periodic motion.
   Amplitide.
   Isolation.
   Superposition.
   Frequency.
   Period.

3. A rock crusher has an acceleration of 5.5 cm. It is spring mounted on a wooden floor, which has a measured acceleration of 3 cm. What is the value for transmission of the vibrations?

4. Find the period of a mechanical vibration if the structure vibrating completes 3000 oscillations in 6 seconds.

5. What is the value for external transmission of vibration if the passenger in a car accelerates 0.1 cm/s^2 while the wheels of the car accelerate 1.6 cm/s^2?

6. If springs in a system similar to that in Example A reduce vibrations to 6% of their original value, what is the acceleration of the support if the acceleration of the machinery is 0.6 m/s^2?

7. Find the acceleration of a support if $T_I = 0.2$ and $a_m = 1 \text{ cm/s}^2$.

11M4

RESONANCE

Each oscillating energy system has one (or more) frequency at which it exhibits its strongest oscillations, called the resonant or natural frequency or frequencies of the system. Energy is most effectively added to such a system when the frequency of the energy source matches a resonant frequency of the system. This module discusses resonance (frequency matching), and gives examples of conditions in which resonance is either an advantage or a disadvantage. Experiments that involve two mechanical systems are used to demonstrate resonance.

OBJECTIVES

To complete this module, the student should be able to:

- Define:

    Natural frequency.
    Resonant frequency.
    Sympathetic vibration.
    Driving frequency.

- Describe the phenomenon of energy transfer (from a driving to a secondary system) when resonance is achieved.
- Describe two cases in which resonance is an advantage, and two cases in which resonance is a disadvantage.
- Calculate the natural frequency of a pendulum, given its length; observe energy transfer from one pendulum to another when their natural frequencies are equal.
- Calculate the resonance frequencies of a closed column of air, given the speed of sound and the length of the air column; observe transfer of acoustic energy from a tuning fork to an air column.

DISCUSSION

All oscillating energy systems have certain frequencies at which they oscillate easily. These frequencies - the ones for which standing waves may be formed - are called the natural frequencies of the system. Because they also are the frequencies at which resonance occurs, they sometimes are referred to as the resonant frequencies.

Resonance is an energy transfer process in which a large amplitude oscillation is established in a system at one of its natural frequencies by driving it with a small amplitude oscillation at that frequency for a period of time. The applied signal is called the driving frequency; the resulting oscillations are termed sympathetic vibrations.

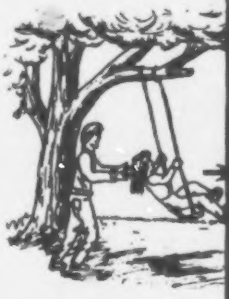


FIG. 1

Figure 1 shows a simple and familiar example of resonance. The driving force in this example is the push given to the swing after it comes to rest at its maximum height on each return motion. The sympathetic vibrations appear as increased amplitude of the swing-like motion. A single push on the swing will have little effect on this amplitude, but a series of small pushes at the proper time will cause a gradual increase in the maximum amplitude.

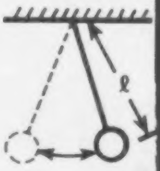


FIG. 2

The period of a pendulum is the time required for it to complete one cycle of its repetitive motion. In Figure 2 this period is the time for the pendulum to swing from its original position to its maximum height in the opposite direction (dotted) and back to the original position. The period of a pendulum is given in Equation 1.

$$T = 2\pi \sqrt{\ell/g} \qquad (1)$$

where

T = period of pendulum.
ℓ = length of pendulum.
g = acceleration of gravity ($9.8\ m/s^2$).

Example A shows how this equation is used in solving a problem.

EXAMPLE A: PERIOD OF A PENDULUM

Given: Pendulum length: 3 m.

Find: Pendulum period.

Solution:

$$T = 2\pi \sqrt{\ell/g}$$

$$= 2 \times 3.14 \times \sqrt{3\text{m}/9.8\ (\text{m/s}^2)}$$

$$= 6.28 \times \sqrt{0.306\ \text{s}^2}$$

$$= 6.28 \times 0.55\ \text{s}$$

$$T = 3.48\ \text{s}$$

The swing of this pendulum will increase in amplitude if it is given a push every 3.48 seconds, which is its natural or resonant frequency. Any other driving frequency will not produce the maximum amplitude.

The closed organ pipe shown in Figure 3 is another example of resonance. Air rushing past the baffle produces a hissing sound that contains many frequencies, including resonant frequencies of the closed pipe; these resonant frequencies produce standing acoustic waves in the pipe. The amplitudes of these waves increase to produce the characteristic sounds of the organ pipe.

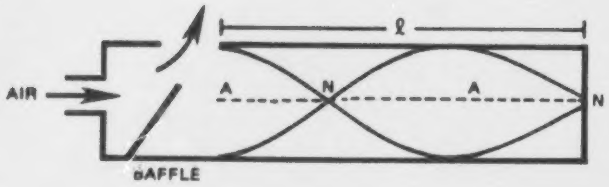


FIG. 3

Other sound frequencies do not produce standing waves in the pipe, and their amplitudes remain too low to be heard. The forward and reflected waves at these frequencies effectively cancel one another out.

The resonant frequencies of the pipe are those that produce a node at the closed end and an antinode at the opening. They are given by Equation 2.

$$f_m = \frac{mv}{4\ell} \qquad (2)$$

where

m = odd integer (1, 3, 5, ...)
v = velocity of sound.
ℓ = length of pipe.

This type of resonant energy system has a large number of resonant frequencies. The lowest of these, corresponding to $m = 1$, is called the "fundamental" frequency. Example B shows the use of Equation 2 in calculating some of these frequencies.

EXAMPLE B: RESONANT FREQUENCIES OF A CLOSED ORGAN PIPE

Given: Closed organ pipe length: 2m; velocity of sound: $v = 346$ m/s.

Find: First three resonant frequencies.

Solution:

$$f_m = \frac{mv}{4\ell}$$

a.

$$f_1 = \frac{1 \times (346 \text{ m/s})}{4 \times (2\text{m})}$$

$$= \frac{346 \text{ m/s}}{8 \text{ m}}$$

$$= 43.25/\text{s}$$

$$f_1 = 43.25 \text{ Hz}$$

b.

$$f_3 = \frac{3 \times (346 \text{ m/s})}{4 \times (2\text{m})}$$

$$f_3 = 129.75 \text{ Hz.}$$

c.

$$f_5 = \frac{5 \times (346 \text{ m/s})}{4 \times (2\text{m})}$$

$$f_5 = 216.25 \text{ Hz}$$

Note that the higher frequencies also can be found by using $f_m = mf_1$.

Resonance, a common phenomen, is used in a variety of applications. The sounds of many musical instruments are produced by resonance in enclosed volumes of air. The ear responds to the sounds because tiny fibers in the inner ear are caused to vibrate at their resonant frequencies.

Electronic circuits containing inductors and capacitors also have resonant frequencies. Electronic signal generators are devices that cause standing electronic waves to be established at these natural frequencies.

Tuners in radios are LC circuits with variable capacitors. When the circuit is adjusted to resonate at the same frequency as a radio station, the amplitude of the signal in the circuit increases to a point at which it may be amplified and converted to sound. Stations on other frequencies do not produce resonance and are not heard.

There are such cases in which resonance must be avoided. If a structure is subjected to a vibration at one of its resonant frequencies, damaging oscillations may result. The most striking example of this was the collapse of the Tacoma Narrows Bridge. Wind exerted a pulsating force on the bridge at one of its natural frequencies. This resonant condition allowed transfer of energy from the wind to the bridge and the amplitude of the vibrations increased, which lead to the collapse of the structure after only a few hours. Since that incident, suspension bridges have been designed to have resonant frequencies far different from the frequencies of natural winds.

Another example of unwanted resonance is the oscillation produced when the wheels of an automobile are not properly balanced. The suspension system of the automobile has a natural frequency at which it will oscillate. At a particular speed the frequency of the vibrations produced by an unbalanced wheel are equal to the resonant frequency of the suspension system; the resulting sympathetic vibrations cause a bumpy ride. The same phenomena can occur in shock mounts for machinery if the vibration frequency of the machinery is near the resonance frequency of the support structure.

Other common examples of resonance include dishes or windows that vibrate when a train or a large truck passes. The vehicle produces sound containing many frequencies. Sympathetic vibrations occur in objects at their natural frequencies if those frequencies are present in the incident sound waves.

SUMMARY

Many mechanical, electrical and fluid systems have natural frequencies at which they oscillate easily. Resonance occurs whenever pulsating energy is delivered to these systems at their natural (resonant) frequency. Some of the incident energy is converted to a wave motion in the system called a sympathetic vibration. In some cases, such as radio tuners and clocks, these resonant frequencies are desirable and useful. In other, such as mechanical structures and d.c. power supplies, they are troublesome and must be eliminated.

RESONANCE OF PENDULUMS

MATERIALS

String (approximately 5m).
Four 100-gram weights.

PROCEDURE

1. Tie length of string between two vertical supports, as shown in Figure 4.
2. Tie a piece of string to each weight and to horizontal string to form four pendulums with lengths indicated in Figure 4.
3. Start right pendulum of Figure 4 swinging in plane perpendicular to horizontal string; observe motion of all four pendulums for several minutes and describe in Data Section.
4. Record length of each pendulum in Data Table 1; calculate period of each (use Equation 1).
5. Explain in a short paragraph how this experiment demonstrates resonance.

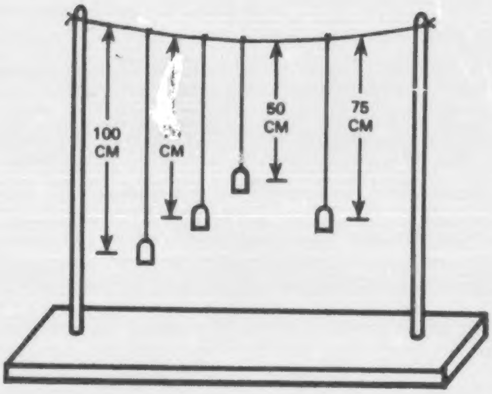


FIG. 4

RESONANCE OF ORGAN PIPES

MATERIALS

Resonance tube (Sargent-Welch #3308) consisting of the following components:

Glass tube 1-inch in diameter and 4-feet long.
One hole stopper for tube.

  - Flexible tubing.
  - Container for water.
  - Meter stick.
  - Mounting brackets.
- Centigrade thermometer.
- Three tuning forks with the following frequencies:
  - 128 Hz.
  - 256 Hz.
  - 512 Hz.

PROCEDURE

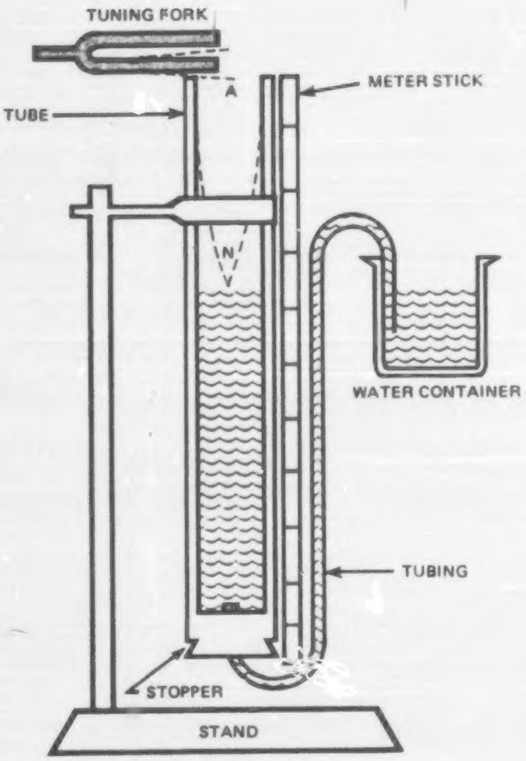


FIG. 5

1. Measure air temperature; record in Data Section.
2. Calculate velocity of sound in air. See equation in Data Section.
3. Calculate length of tube required for resonance at frequency of each tuning fork (use Equation 2); record lengths in Data Section.
4. Set up equipment as shown in Figure 5.
5. Fill glass tube and flexible tube with water; adjust water level in glass tube by raising or lowering water container until length of tube containing air is 2 or 3 cm less than calculated length for resonance for lowest frequency tuning fork.
6. Strike lowest frequency fork and hold over tube end, as shown in Figure 5; simultaneously, lower water container to allow water to drain from tube; listen to tone produced. (Resonance, which will occur when correct air length is reached, is indicated by maximum amplitude of sound.) Repeat this step with water level rising through resonance point; record average resonant length in Data Table 2.
7. Repeat steps 5 and 6 for other two tuning forks.
8. In a brief paragraph, explain what occurs in the air column at resonance.

RESONANCE OF PENDULUMS

Describe the motions of each of the four pendulums during the experiment performed in the Laboratory procedures:

DATA TABLE 1.

Pendulum	Length (m)	Period (s)
1		
2		
3		
4		

RESONANCE IN ORGAN PIPES

Air Temperature: T = ______°C

Velocity of Sound in Air: V = 331 m/s + (0.6 m/s - C°) x t
= 331 m/s + (0.6 m/s - C°) x (____C°)
= _______ m/s.

DATA TABLE 2.		
Frequency of tuning fork (Hz)	Calculated Length of air column for first resonance point (m) $\ell = \frac{v}{4f}$	Measured Length of air column for resonance (m)

SELF CHECK

1. Identify the terms that match the following definitions:

   a. The frequency at which an oscillating energy system vibrates most easily (two names).
   b. The vibrations occuring in a system when it receives energy at one of the above frequencies.
   c. The frequency of the energy applied to a resonant system.

2. Describe how energy is transferred to a pendulum when a force is applied at its resonance frequency.

3. Describe how resonance is used advantageously in a radio tuner.

4. Describe how resonance can be a disadvantage in structures (such as bridges).

5. Calculate the period of pendulums with each of the following lengths:

   a. 25 cm.
   b. 5 m.

6. Calculate the fundamental frequency of a closed organ pipe with each of the following lengths:

   a. 50 cm.
   b. 37.5 cm.

7. What is the frequency of a simple pendulum if the pendulum is 2 meters long?

8. What is the length of a closed organ pipe if the fifth resonant frequency is 270 Hz? Use 350 m/s as the speed of sound.

9. What is the length of a pendulum whose period is 1 second?

11F1

CAVITY VIBRATIONS AND STANDING WAVES

Many devices require, to some extent, the formation of standing waves to be able to perform their functions. These devices include musical instruments, microwave generators and lasers. Standing waves are produced when two like waves of equal amplitude and frequency interfere with each other while traveling in opposite directions. A cavity is a volume containing a medium that supports wave reflection at the ends of the volume. This module investigates the nature of standing waves and conditions necessary for their formation. In the laboratory, the student will produce acoustic standing waves in an "organ pipe" and determine characteristics of cavity vibrations.

OBJECTIVES

Upon completion of this module, the student should be able to:

- Given any two of the following quantities, calculate the third:

    Wavelength.
    Frequency.
    Wave velocity.

- Define the term "standing wave."
- Given the wave velocity and length of a resonant cavity, calculate resonant frequencies.
- Given appropriate equipment, construct a resonant acoustic cavity and determine its resonant frequencies.

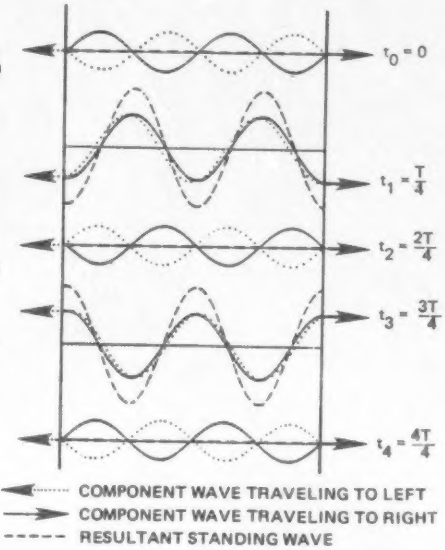
382

DISCUSSION

A standing wave is formed when two like waves of equal amplitude and frequency interfere with one another while traveling in opposite directions. The waves may be transverse waves, such as waves in a string, or longitudinal waves, such as sound waves. The resultant is a wave that appears to "stand still" in space. That is, certain regions exist wherein displacement is quite large, while other regions exist wherein there is little or no displacement. Figure 1 illustrates how this occurs and why the wave appears to stand still.

In this figure, two waves (dotted and solid lines) of equal amplitude and frequency travel in opposite directions. The waves are exactly out of phase at t_0, which produces a resultant represented by a straight line. At t_1 (1/4 the wave period later), the solid wave has moved 1/4 a wavelength to the right, and the dotted wave has moved 1/4 a wavelength to the left. This movement places the two waves in phase and gives the maximum resultant shown. At t_2 the waves are out of phase again. At t_3 they are in phase again to produce a maximum displacement. At t_4 the period of each wave is completed and the process repeats.

As time proceeds, the standing wave shows points of maximum vibration (antinodes) separated by points of zero vibration (nodes) (A and N in Figure 2). At antinodal positions (A), the resultant wave moves up and down between maximum displacement limits shown (arrows in Figure 2). At nodal positions (N), no up and down motion occurs. Overall, there is no apparent motion of the resultant wave in a left-right direction. Hence, the resultant wave appears to stand still even though individual waves are moving to the left and right.



T IS THE PERIOD OF WAVE MOTIONS

FIG. 1

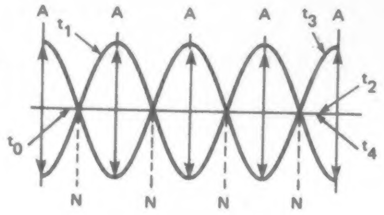


FIG. 2

Figure 3 shows several possible standing waves in a string tied at both ends. Since the string cannot vibrate at the ends, there must always be a node there. Wave velocity remains constant, depending only on the tension in the string and its density. Wavelength and frequency of each pattern or mode are given in the figure. Notice that relationship between wavelength and string length can be found by inspection of the figure. Once the wavelength for any mode is determined, corresponding frequency

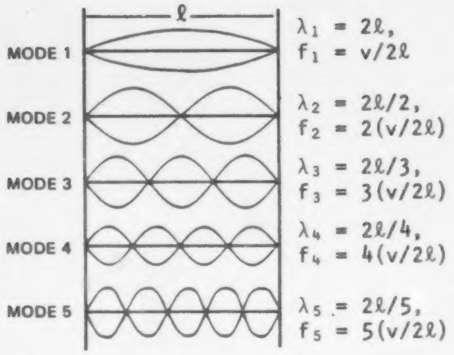

ℓ
MODE 1
$\lambda_1 = 2\ell$,
$f_1 = v/2\ell$
MODE 2
$\lambda_2 = 2\ell/2$,
$f_2 = 2(v/2\ell)$
MODE 3
$\lambda_3 = 2\ell/3$,
$f_3 = 3(v/2\ell)$
MODE 4
$\lambda_4 = 2\ell/4$,
$f_4 = 4(v/2\ell)$
MODE 5
$\lambda_5 = 2\ell/5$,
$f_5 = 5(v/2\ell)$

FIG. 3

can be found from Equation 1.

$$f = \frac{v}{\lambda} \tag{1}$$

where
f = frequency.
v = wave velocity.
λ = wavelength.

Notice that mode spacing - the frequency difference between two adjacent modes - is a constant given by Equation 2.

$$\Delta f = \frac{v}{2\ell} \tag{2}$$

where
Δf = mode spacing.
ℓ = cavity length.

Frequency of any mode can be determined by multiplying this frequency by the mode number, as given in Equation 3:

$$f_n = n\Delta f \tag{3}$$

where
f_n = frequency of nth mode.
n = mode number.

Example A shows use of Equations 1-3 in solving a problem.

Resonant frequencies of a system are those frequencies for which a standing wave can exist. In Example A, any multiple of 500 Hz is a resonant frequency. No other frequency can form a standing wave since it would not produce a node at both ends of the string.

Figure 4 shows the first four modes possible in an organ pipe with both ends open. In this case, the air executing longitudinal vibrations is free to move at each end. This results in an antinode at each end of the pipe. Equations developed from Figure 3 hold for this example also.

A "cavity" is any volume containing a medium that supports wave propagation in such a manner that some reflection of the waves occurs at the ends of the volume. In the case of the vibrating string, relflections occur at the tied ends. In the case of the

EXAMPLE A: CALCULATION OF MODE SPACING, FREQUENCY AND WAVELENGTH

Given: From Figure 3, $\ell = 1$ m and $v = 1000$ m/s

Find: (a) mode spacing, (b) frequency of third mode and (c) wavelength of third mode.

Solution: (a)
$$\Delta f = \frac{v}{2\ell} = \frac{1000 \text{ m/s}}{2 \times (1 \text{ m})}$$
$$\Delta f = 500 \text{ Hz}$$

(b)
$$f_n = n\Delta f$$
$$f_3 = 3 \times 500 \text{ Hz} = 1500 \text{ Hz}$$

(c)
$$\lambda = \frac{v}{f} = \frac{1000 \text{ m/s}}{1500 \text{ Hz}} = 0.67 \text{ m}$$

Notice that in Figure 3, wavelength of the third mode is 2/3 ℓ, or 0.67 m, as determined in this example.

organ pipes, air transmits sound waves, and open ends of the pipes actually produce partial reflection of the waves. In the optical cavity of a laser, light is reflected by a mirror at each end. In each case, reflection of waves at the ends of the cavity provides a pair of identical waves traveling in opposite directions. If frequency of these waves is a resonant frequency of the cavity, a standing wave is formed.

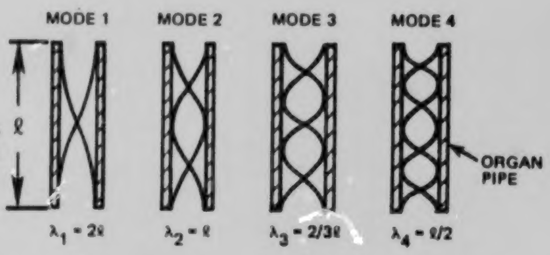


FIG. 4

Resonant cavities are employed when a pure frequency is to be produced. Most musical instruments use resonant cavities consisting of columns of air (organ pipes, trumpets) or stretched strings (pianos, violins). Klystron tubes use a metallic-walled cavity to produce standing waves in the microwave region of the electromagnetic spectrum. Such tubes are employed in radar and communication.

Figure 5 shows elements of a klystron tube. Electrons from the cathode are accelerated toward the anode. They pass through the grids into the drift space. They are reflected back through the upper grid by the repeller. Velocity changes of the electrons passing through the cavity produce microwave radiation. The cavity of the tube is a resonant cavity for the wavelength of microwaves to be produced, and standing waves form within it. Part of the energy is transmitted out of the tube by the output coupling loop. Wavelength of the microwaves produced can be altered by changing the shape of the resonant cavity.

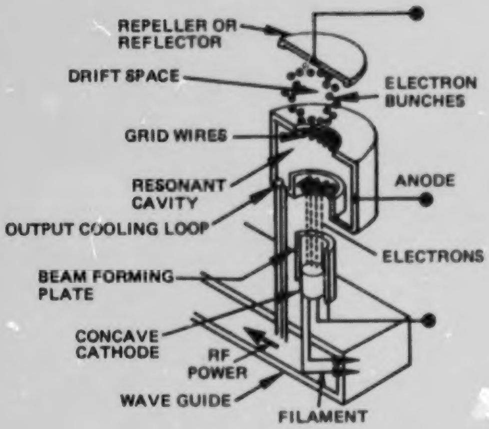

REPELLER OR REFLECTOR
DRIFT SPACE
ELECTRON BUNCHES
GRID WIRES
RESONANT CAVITY
ANODE
OUTPUT COOLING LOOP
ELECTRONS
BEAM FORMING PLATE
CONCAVE CATHODE
RF POWER
WAVE GUIDE
FILAMENT

FIG. 5

SUMMARY

Standing waves result from the interference (constructive or destructive) produced by the superposition of waves traveling in opposite directions. Many lasers, microwave generators, musical instruments and other devices require some formation of standing waves. Waves in a cavity reflect at the ends of the cavity to produce a pair of identical waves traveling in opposite directions; if these wave pairs are in resonance, they are standing waves.

STANDING WAVES AND CHARACTERISTICS OF CAVITY VIBRATIONS

MATERIALS

Oscilloscope.
Audio frequency signal generator.
Speaker.
Microphone.
Cardboard or plastic tube approximately 3 inches in diameter and 4 feet long, open at both ends.
Similar tube approximately 2 feet long.
Clip leads.

PROCEDURE

1. Measure length of each tube; record in Data Table.

2. Set up equipment as shown in Figure 6, using longer of two tubes. Set oscilloscope for 0.005 v/cm and 2 ms/cm. Turn on signal generator; set frequency control at approximately 200 Hz.

3. Slowly increase frequency produced by signal generator while watching display on oscilloscope. Amplitude of displayed audio signal will increase and decrease as frequency increases. Maximum amplitude occurs when frequency of sound waves is a resonant frequency of open tube and acoustic standing waves are formed in tube. Measure at least four of these frequencies with oscilloscope; record them in order in Data Table [lowest frequency (f_1) first, and so on up to f_4].

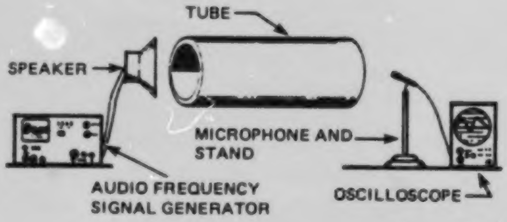


FIG 6

4. Complete calculations in part A of Data Table.

5. Complete predicted results portion of part B of Data Table, using velocity of sound in air determined in part A.

6. Replace original tube with shorter one; determine its resonant frequencies. Record these in experimental-results

section of Data Table.

7. Compare predicted and experimental resonant frequencies in part B of Data Table.

8. How does the cavity size formed by the tubes affect wave resonance?

DATA TABLE

PART A

Length of first tube: ℓ = ________m

Resonant frequencies: f_1 = ________Hz
f_2 = ________Hz
f_3 = ________Hz
f_4 = ________Hz

Mode spacings: $f_4 - f_3$ = ________Hz
$f_3 - f_2$ = ________Hz
$f_2 - f_1$ = ________Hz

Average modue spacing:

$$\Delta f = \frac{1}{3}[(f_4 - f_3) + (f_3 - f_2) + (f_2 - f_1)]$$

$$= \frac{1}{3}[\_\_\_\_\_\_\_\_\text{Hz} + \_\_\_\_\_\_\_\_\text{Hz} + \_\_\_\_\_\_\_\_\text{Hz}]$$

Δf = ________Hz

Since $\Delta f = \frac{v}{2\ell}$,

velocity of sound: $v = \Delta f \times 2\ell$
= ________Hz x 2 x ________m
v = ________m/s

PART B

Length of second tube = ________m

Predicted results:

Mode spacing: $\Delta f = \frac{v}{2\ell}$
= ________m/s ÷ (2 x ________m)
Δf = ________Hz

Resonant frequencies: $f_1 = 1 \times \Delta f$ = ________Hz
$f_2 = 2 \times \Delta f$ = ________Hz
$f_3 = 3 \times \Delta f$ = ________Hz
$f_4 = 4 \times \Delta f$ = ________Hz

Experimental Results:

Measured resonant frequencies: f_1 = ________ Hz
f_2 = ________ Hz
f_3 = ________ Hz
f_4 = ________ Hz

SELF CHECK

1. Velocity of light in vacuum is 3×10^8 m/s. What is the frequency of a light wave with a wavelength of 5×10^{-5} cm?
2. Velocity of sound in air at sea level is 335 m/s. What is wavelength of a sound wave with a frequency of 800 Hz?
3. Define the term "standing wave."
4. An organ pipe (open at both ends) has a length of 3.2 m. What are frequencies of the first three modes of vibration?
5. A laser cavity with mirrors at opposite ends has a length of 25 cm. Calculate axial mode spacing in the cavity. Use 3×10^8 m/s as velocity of light in the cavity.
6. Predict the first few modes of vibration for an organ pipe closed at one end and open at the other. Give mode frequencies in terms of pipe length ℓ and velocity of sound v in air.

11F4

DOPPLER EFFECT

The pitch of the whistle of a locomotive sounds higher when it is approaching then when it is receding from a stationary listener. This phenomenon was explained by an Austrian named Christain Johann Doppler in 1842 and applies, not only to sound waves, but to waves in general. For example, the color of a luminous body must change due to its motion relative to an observer.

The Doppler effect has proved to be, not only an interesting physical phenomenon, but also a useful one. It has enabled astronomers to determine the relative speeds of galaxies, enriching their insight into the origin and future of the universe. Almost everyone is familiar with the defense and traffic-control applications of radar, which operates on the Doppler principle. The production of shock waves by supersonic aircraft is another common occurrence accompanying the Doppler effect.

In this module, the Doppler effect is presented for both sound and electromagnetic waves, including appropriate equations for determining the frequency shifts that result. The phenomenon of shock waves, and applications to astronomy, spectroscopy, and radar are discussed. In the laboratory, the student will measure the Doppler frequency of sound waves.

OBJECTIVES

Upon completion of this module, the student should be able to:

- Calculate the Doppler frequency of a sound wave, given speeds of the source and observer and the frequency if both are at rest.

- Explain how beats are produced; define "beat frequency."
- Draw a diagram of a shock wave, and explain how it is produced; define "Mach cone" and "Mach number"; calculate the angle included in the Mach cone.
- Calculate the Doppler frequency of electromagnetic radiation, given the frequency of the source and its relative speed.
- Explain how radar is used to track a moving object, and determine its speed; calculate the velocity of a moving object relative to a radar transmitter, given the frequency and Doppler frequency.
- Explain the Doppler broadening of absorption and emission spectral lines.
- Calculate the relative velocity of a luminous heavenly body, given the Doppler shift of one of its spectral lines.
- Given the appropriate equipment, measure the Doppler shift of sound waves.

DISCUSSION

Figure 1 depicts sound waves produced by a stationary source S and perceived by a stationary observer at O. The concentric circles represent wave fronts. The frequency of the sound heard by the observer is the same as that heard if he or she were positioned at the source. Wave frequency is indicated by Equation 1.

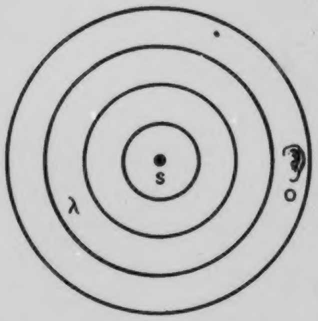


FIG. 1

$$f = \frac{v}{\lambda} \qquad (1)$$

where

f = frequency of source.
v = velocity of waves.
λ = wavelength.

On the other hand, if the observer moves toward the source at velocity v_o as shown in Figure 2, he encounters each wavefront sooner than it normally would reach him if he were stationary. The wavefronts perceived by the moving observer appear to be more closely spaced, thereby encountered at a higher frequency. The observer, therefore, hears a higher pitch than when he is stationary.

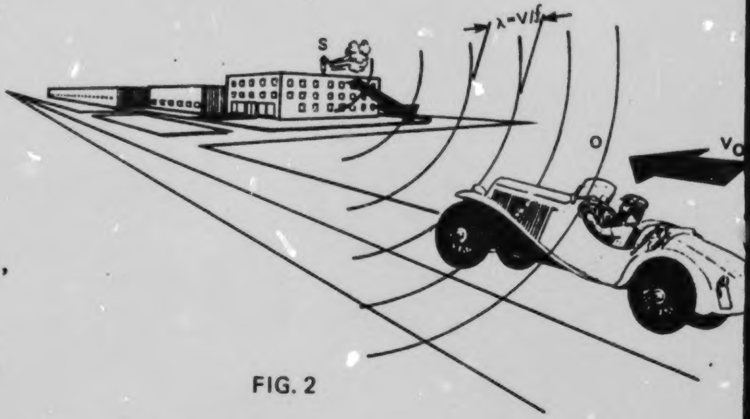


FIG. 2

The number of waves received by the stationary observer in time t is vt/λ. When he moves toward the source, he encounters the waves at a higher rate. He must receive an additional $v_o t/\lambda$ waves in the same time interval t. The frequency f_o heard by the observer is the number of waves received in time t divided by t, as in Equation 2.

$$f_o = \frac{\frac{vt}{\lambda} + \frac{v_o t}{\lambda}}{t}$$

$$= \frac{v}{\lambda} + \frac{v_o}{\lambda} \qquad \text{(Dividing by t)}$$

$$= \left(\frac{v}{\lambda}\right)\left(1 + \frac{v_o}{v}\right) \qquad \text{(Factoring)}$$

$$f_o = f_s\left(1 + \frac{v_o}{v}\right) \qquad \text{(Since } f = v/\lambda\text{)} \qquad (2)$$

where

f_o = frequency of waves heard by observer.
f_s = frequency of waves at source.
v = velocity of sound in air = 330 m/s.
v_o = velocity of observer.

The frequency f_o heard by the observer is the frequency f_s heard if $v_o = 0$ (observer stationary), plus the increase fv_o/v due to the motion of the observer. The frequency f_o is called the "Doppler frequency."

If the observer is moving away from the source, he receives $v_o t/\lambda$ fewer waves in time t. The Doppler frequency is then given by Equation 3.

$$f_o = f_s\left(1 - \frac{v_o}{v}\right) \tag{3}$$

The listener moving away from the source hears a lower pitch, or frequency, than he would if he were stationary. Example A illustrates how the "Doppler shift" — or change in frequency, $(f_o - f_s) = (\pm f_s\ v_o/v)$ — is found for a moving observer.

EXAMPLE A: DOPPLER EFFECT FOR A MOVING OBSERVER

Given: An automobile traveling at 27 m/s approaches a factory sounding its whistle at a frequency of 1980 Hz.

Find: The frequency heard by the motorist (a) as he approaches and (b) after he passes the factory and moves away.

Solution: The velocity v_o of the observer is 27 m/s, and that of the wave is 330 m/s, the velocity of sound in air. On approach, the driver hears —

$$\begin{aligned} f_o &= f_s\left(1 + \frac{v_o}{v}\right) \\ &= (1980\ \text{Hz})\left(1 + \frac{27}{330}\right) \\ f_o &= 2142\ \text{Hz} \end{aligned}$$

As he moves away he hears —

$$\begin{aligned} f_o &= f_s\left(1 - \frac{v_o}{v}\right) \\ &= (1980\ \text{Hz})\left(1 - \frac{27}{330}\right) \\ f_o &= 1818\ \text{Hz} \end{aligned}$$

The Doppler shift in either case is $f_o - f_s = \pm 162$ Hz

In the case of the moving observer and stationary source, the actual wavelength $\lambda = v/f$ of the sound waves does not change. If, however, the observer remains stationary and the source moves toward him at a speed v_s as in Figure 3, the wave fronts received by the observer are closer together. This effect occurs because with each new wave emitted by the source, the source is closer to the observer. The wavelength is, therefore, shortened by the distance v_s/f_s traveled by the source in each cycle. The Doppler frequency f_o is found in Equation 4 developed below.

$$
\begin{aligned}
f_o &= \frac{v}{\lambda_o} \\
&= \frac{v}{\left(\lambda - \frac{v_s}{f_s}\right)} && \text{(Since } \lambda_o = \lambda - v_s/f_s\text{)} \\
&= \frac{f_s v}{\lambda f_s - v_s} && \text{(Simplifying the fraction)} \\
f_o &= \frac{f_s v}{(v - v_s)} && \text{(Since } v = \lambda f_s\text{)} \qquad (4)
\end{aligned}
$$

where

λ_o = wavelength received by observer as source approaches.

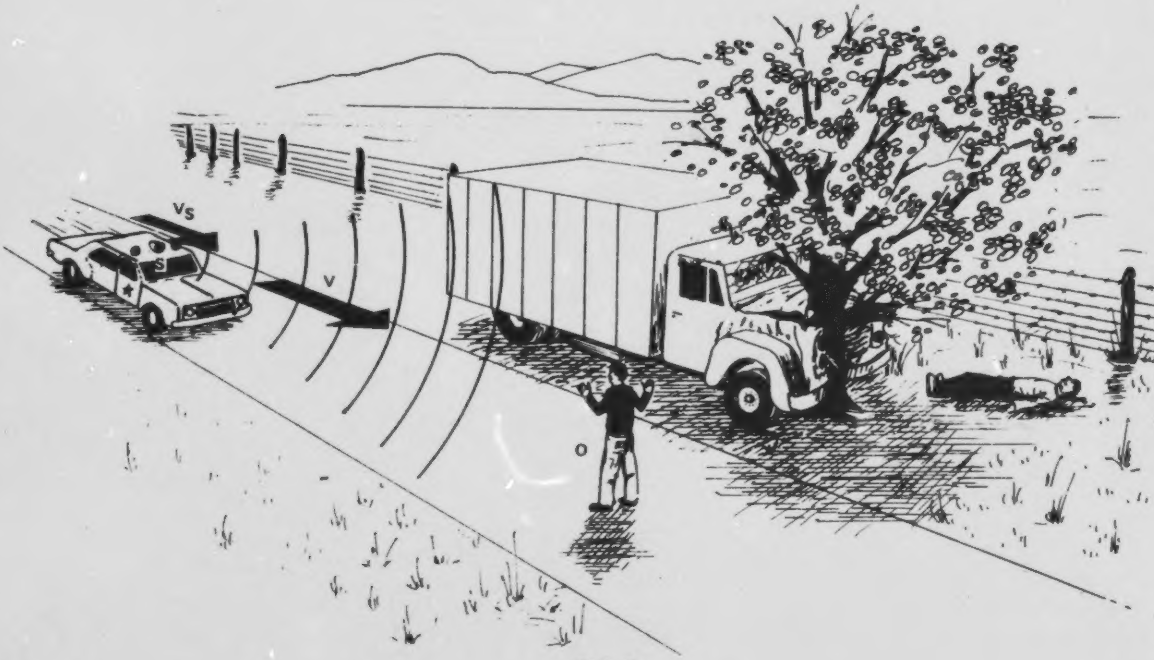


FIG. 3

Since the waves received are closer together, the frequency heard by the observer is higher than if the source were stationary. When the source moves away from the observer, the wavelength is now larger by v_s/f, and the Doppler frequency becomes that indicated by Equation 5. In this equation, f_o is less or lower in pitch than f_s.

$$f_o = \frac{f_s v}{(v + v_s)} \qquad (5)$$

The Doppler effect for sound was first tested experimentally in 1845 by Buys Ballot. A trumpet player blew his horn while riding on a flatcar pulled by a locomotive, as in Figure 4a. A second trumpet player stood beside the tracks and blew his horn simultaneously at the same frequency. The sound waves from the two trumpets are illustrated in Figure 4b. Notice that the frequency of the moving trumpet is increased slightly as it approaches the stationary player. The waves are shown plotted together in Figure 4c. Since the waves interfere, they can be added together to give a resultant wave like the one shown in Figure 4d. The maxima of such a resultant wave are called "beats." The beat frequency is the difference between the frequencies of the two interfering waves, and is the Doppler shift in frequency.

Examples B and

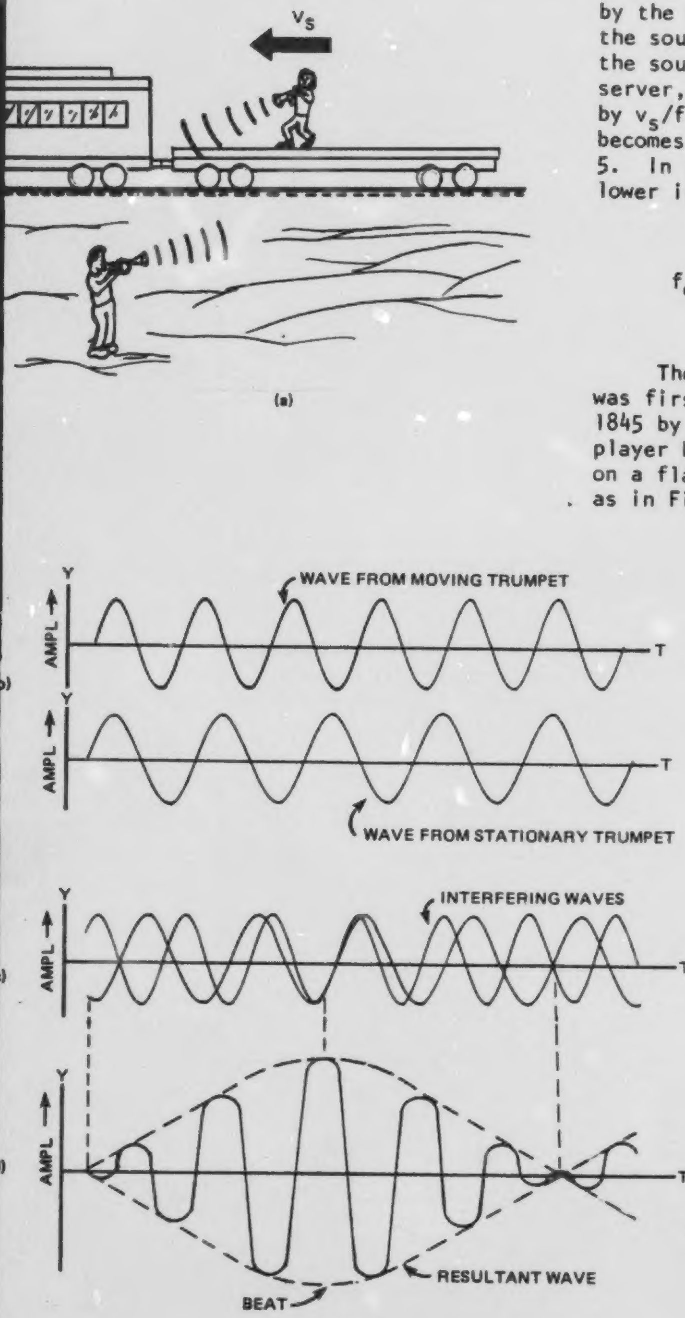


FIG. 4

C illustrate the Doppler shift for a moving source. In both examples the velocity (v) of sourced through air is 330 m/s.

EXAMPLE B: DOPPLER EFFECT FOR A MOVING SOURCE

Given: A trumpet player blowing his horn at 440 Hz approaches at a speed of 3 m/s a stationary player blowing his horn at the same frequency.

Find: The Doppler shift in frequency.

Solution: The speed of the source v_s is 3 m/s.

$$f_o = \frac{f_s v}{(v - v_s)}$$

$$= \frac{(440\ \text{Hz})(330)}{(330 - 3)}$$

$$f_o = 444\ \text{Hz}$$

The Doppler shift or beat frequency is —

$$f_o - f_s = 444 - 440 = 4\ \text{Hz}$$

EXAMPLE C: DOPPLER EFFECT FOR A MOVING SOURCE

Given: A siren emitting sound of frequency 1000 Hz moves away from a stationary observer at 40 m/s.

Find: The frequency heard by the observer.

Solution: For receding sources, the pitch or frequency is lower.

$$f_o = \frac{f_s v}{(v + v_s)}$$

$$= \frac{(1000\ \text{Hz})(330)}{330 + 40}$$

$$f_o = 892\ \text{Hz}$$

The mathematical results of the Doppler effect for sound can be summerized by Equation 6 for all cases considered.

$$f_o = \frac{f_s(v \pm v_o)}{(v \pm v_s)} \qquad (6)$$

Equation 6 allows treatment of cases in which both the source and the observer are in motion. The appropriate choice of sign must be memorized because the frequency f_o must be higher than f_s for the converging source and observer and must be lower for the receding source and observer. Cases in which both the source and the observer are moving are presented in Examples D and E.

EXAMPLE D: DOPPLER SHIFT FOR BOTH OBSERVER AND SOURCE MOVING

Given: An automobile traveling at 20 m/s due east approaches another traveling at 25 m/s due west. The westbound driver sounds his horn, which has a frequency of 1200 Hz.

Find: The frequency heard by the eastbound motorist.

Solution: The Doppler frequency is higher than 1200 Hz since the cars are approaching each other. Thus, one chooses the positive sign for the numerator and the negative sign for the denominator.

$$f_o = \frac{f_s\ (v + v_o)}{(v - v_s)}$$

$$= \frac{(1200\ \text{Hz})(330 + 20)}{(330 - 25)}$$

$$f_o = 1377\ \text{Hz}$$

EXAMPLE E: DOPPLER SHIFT FOR BOTH OBSERVER AND SOURCE MOVING

Given: A car with a siren of frequency 2000 Hz traveling at 30 m/s is chasing an observer who is moving at 35 m/s.

Find: The frequency of the siren heard by the observer.

Solution: The observer moves away from the source, and the source moves toward the observer.

$$f_o = \frac{f_s\ (v - v_o)}{(v - v_s)}$$

$$= \frac{(2000\ \text{Hz})(330 - 35)}{(330 - 30)}$$

$$f_o = 1967\ \text{Hz}$$

SHOCK WAVES

When the wave source moves through a medium at a faster rate than the speed of the wave in the medium, the wavefront becomes conical in shape, with the moving source at its apex (Figure 5). Examples of this type of motion are the bow wave of a boat and the shock wave from a supersonic aircraft or from a bullet. The waves pile up at the surface of the cone, causing a severe pressure disturbance. The angle θ between the velocity v_s of the source and the surface of the cone is given by Equation 7 and illustrated by Figure 5. Here v is the velocity of the wave disturbance.

$$\sin \theta = \frac{vt}{v_s t} = \frac{v}{v_s} \qquad (7)$$

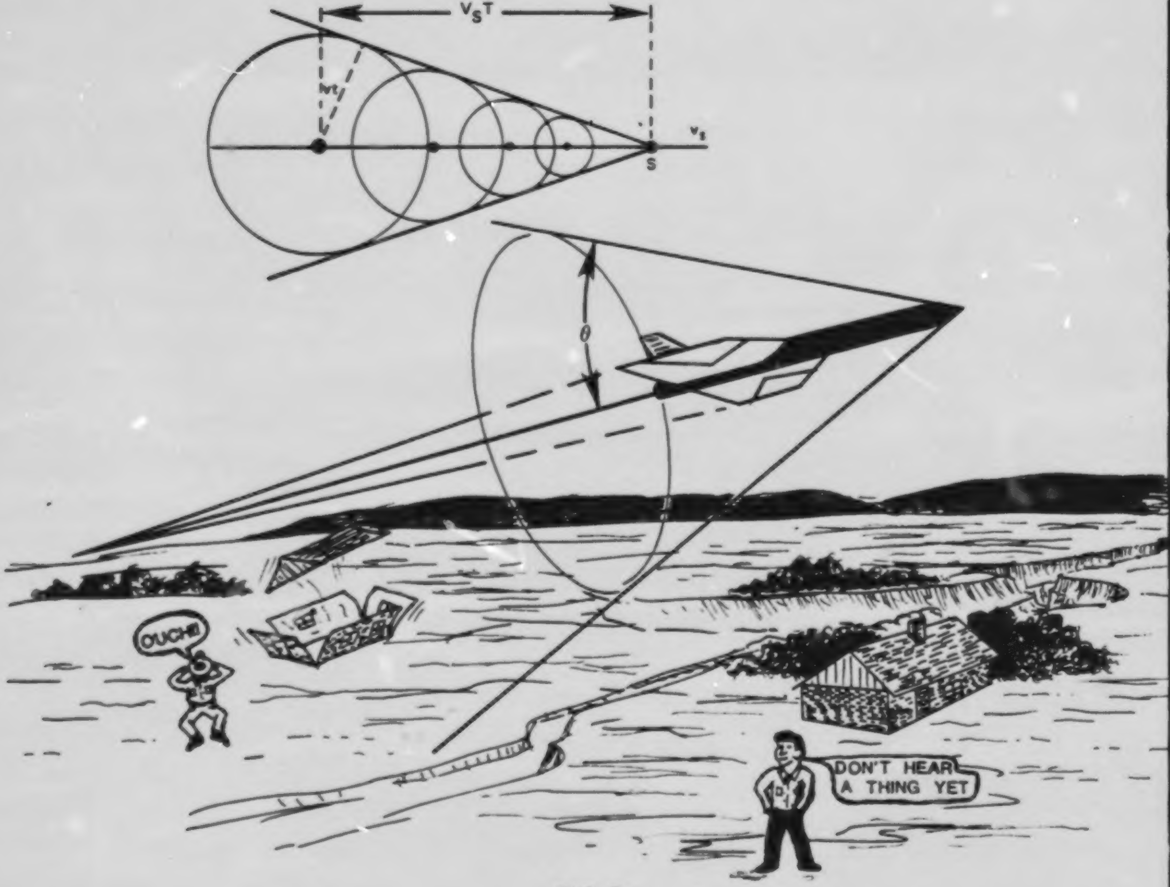


FIG. 5

The cone produced by a supersonic aircraft is called the "Mach cone," and the ratio v_s/v the "Mach number," after the physicist Ernst Mach. When an aircraft "breaks the sound barrier," it is traveling faster than 330 m/s, and the surface of the Mach cone which reaches observers on the ground, causes them to hear a sharp crack. The cone reduces to two intersecting lines for a bow wave.

Example F shows the calculation of the angle of a shock wave.

EXAMPLE F: SHOCK WAVES

Given: A destroyer traveling at 20 knots (1 knot = 1.15 miles/hour) generates waves traveling at 8 knots.

Find: The angle θ between the ship's direction and its bow wave.

Solution: From Equation 7 —

$$\theta = \sin^{-1}\left(\frac{v}{v_s}\right)$$

$$= \sin^{-1}\left(\frac{8}{20}\right)$$

$$\theta = 23.6°$$

THE DOPPLER EFFECT FOR ELECTROMAGNETIC WAVES

One might expect that a moving light source would appear to be changed in color, just as a moving sound source appears to be changed in pitch; however, the equations developed for sound do not apply. The velocities v_o and v_s were defined with respect to a particular medium in which sound waves travel at speed v. Sound, being a pressure wave, requires such a medium for transmission. No such medium of transmission is required for electromagnetic waves, which can travel through a vacuum. This does not mean there is no Doppler shift for light waves; rather, it implies that only the speed of the source relative to the observer can have any significance.

For nonrelativistic speeds ($u \ll c$), the Doppler frequency is calculated by use of Equation 8.

$$f_o = f_s \left(1 \pm \frac{u}{c}\right) \qquad (8)$$

where

u = relative motion of source and observer.
c = speed of light in a vacuum.
f_s = frequency of the source.
f_o = Doppler-shifted frequency as detected by observer.

The choice of sign is determined just as for sound waves.

The Doppler effect for radio waves, which have wavelengths that extend from a few centimeters up to 10^5 meters, is used for radar traffic control and defense systems. Radar permits measurement of the beat frequency between emitted radio waves and the reflected wave from an approaching or receding object to determine its speed (Figure 6). Example G illustrates how radar is used in traffic control.

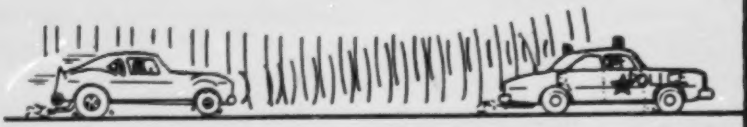


FIG. 6

EXAMPLE G: DOPPLER EFFECT FOR RADIO WAVES

Given: Radio waves having a frequency of 3×10^9 Hz (3GHz) are used to measure the speed of an approaching motorist on the highway. The beat frequency measured is 300 beats/s.

Find: The speed of the approaching automobile.

Solution: The beat frequency is equal to the Doppler shift in frequency, $(f_o - f_s)$.

$f_o = f_s\,(1 + u/c)$ (used + sign since auto is approaching)

$$f_o - f_s = f_s\,(\frac{u}{c}) \qquad \text{(rearranged Equation 8)}$$

$$300 = (3 \times 10^9 \text{ Hz})(\frac{u}{3 \times 10^8 \text{ m/s}})$$

$$u = \frac{(300)(3 \times 10^8 \text{ m/s})}{3 \times 10^9}$$

u = 30 m/s, or about 67 mph

The shift in wavelength that results from the Doppler effect for electromagnetic waves for nonrelativistic relative speeds u is $\lambda_o(\pm u/c)$. Since c is 3×10^8 m/s, a change in color is not observable to a motorist approaching a traffic light; the shift for a typical automobile speed of 20 m/s and light of wavelength 6000 A° is 6000 A° $\times$ 20 / (3×10^8), or 4×10^{-4} A°, which is 7×10^{-6}%!

The Doppler shift in wavelength becomes significant at higher relative speeds u. Gas molecules heated by an electric discharge emit light of wavelengths characteristic of the gas. Because the molecules are very energetic, they have large velocities. Doppler shifts occur in their spectra, the magnitude and direction of the shift depending upon the speed of the individual molecule producing a particular spectral line, and upon whether or not it is moving toward or away from the photographic plate. Similar shifts occur in the absorption spectra from such molecules when white light shines through the gas and when characteristic wavelengths are selectively absorbed. The resulting absorption and emission spectral lines, representing many molecules, are broadened. Figure 7 illustrates a method of viewing such a spectrum for sodium vapor.

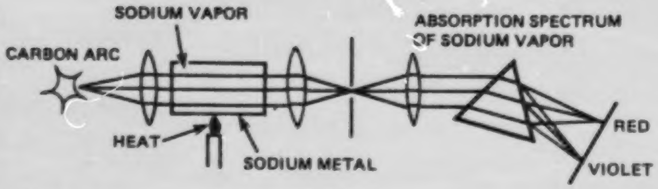


FIG. 7

Absorption spectral lines from a star are not only broadened because of random molecular motions, but also are shifted considerably because of the relative motion of the star with respect to the earth. Such shifts in wavelengths are used to determine the velocities of stars and of entire galaxies relative to the earth. Science has determined that all the distant galaxies that have been studied are moving away from the earth with speeds proportional to their distances; this finding establishes the model of the universe as an expanding one and suggests a primal "big bang" as the origin of the expansion.

Example H consists of the solution to a problem involving the relative speed of a galaxy.

EXAMPLE H: RELATIVE SPEED OF A GALAXY

Given: The Doppler wavelength of a spectral line known to be of wavelength 6560 A° is measured as 6586 A°.

Find: Velocity of the galaxy relative to the earth.

Solution: The Doppler wavelength of the spectral line is increased by 6586 - 6560 = 26 A°. The galaxy, therefore, is receding. The velocity is found as follows:

$$26 \text{ A}^\circ = 6586 \text{ A}^\circ \frac{u}{(3 \times 10^8 \text{ m/s})}$$

$$u = \frac{(26)(3 \times 10^8 \text{ m/s})}{6586} \quad \text{(solving for u)}$$

$$u = 1.2 \times 10^6 \text{ m/s}$$

SUMMARY

The Doppler effect is the apparent change in the frequency of sound or of electromagnetic radiation due to relative motion between the source and the observer. The pitch, or frequency, of the sound emitted by a moving object (e.g. the whistle of a moving train) appears to a stationary observer to increase as the object approaches and to decrease as it recedes. If the source is stationary, an observer moving toward it intercepts more sound waves per unit time than if the observer were stationary; consequently, the observer hears a higher frequency. When the observer moves away from the source, fewer of the waves catch up with the observer during a unit time, and thus, the observer hears a lower frequency.

The Doppler effect is utilized widely in the area of electromagnetic radiation. The light emitted by a moving object appears more red (red light being of lower frequency than the other colors) when it is receding from the observer (or the observer receding from it). Thus, the fact that light emitted by the stars of distant galaxies experiences a shift when observed from the earth is taken to mean that these distant galaxies are receding from our galaxy. This is the principal evidence for the widely-accepted hypothesis concerning the expansion of the universe. The Doppler effect is used also in radar to distinguish between stationary and moving targets and to provide information concerning their velocity.

DOPPLER SHIFT OF SOUND WAVES

MATERIALS

Stroboscope.
Oscilloscope.
Microphone.
Variable speed electric drill
Vise.
One of the following (1 or 2):

1. Bolt and nut.
   Meter stick cut to 30 cm in length.
   Sound source (tuning fork and rubber mallet or oscillator circuit with speaker).
   Counterweight and clamps.
2. Doppler's Principle Apparatus (Sargent Welch 0964) or equivalent and compressed air.

PROCEDURE

If Doppler's Principle Apparatus is available, omit Step 1.

1. Mount the meter stick on the drill, using the nut and bolt in such a manner that the meter stick turns like a propeller. (See Figure 8.) Mount the sound source at one end of the meter stick, balanced by a counterweight at the opposite end if necessary. Be sure that both the source and the counterweight are clamped securely.
2. Determine frequency of stationary sound source as follows: Adjust oscilloscope so that the sound source produces a clear sine wave on the screen (Figure 8). Record the seconds per division (SEC/DIV) and the number of divisions in one cycle in Data Table. The period T of the wave is the number of divisions in one cycle, multiplied by SEC/DIV. Frequency is the inverse of the period T: f = 1/T. Calculate f and record in Data Table

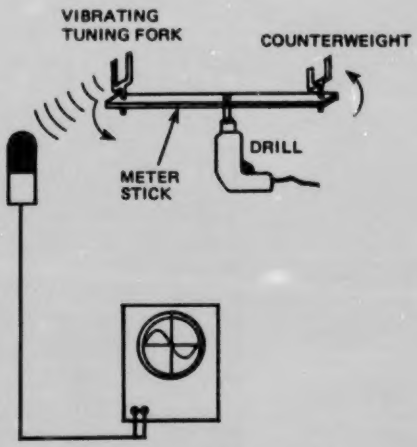


FIG. 8

under RESULTS.

3. With the sound source rotating, determine the minimum and maximum periods, T_{min} and T_{max}, as in procedure 2. Record data and results. The Doppler frequencies for approach and recession are found as follows:

$$f_A = 1/T_{min} \quad \text{(approach)}$$

$$f_R = 1/T_{max} \quad \text{(recession)}$$

Calculate and record these values.

4. Determine velocity v_s of source as follows: Measure and record radius of rotation r in Data Table. Use stroboscope to determine the angular velocity ω of drill, or use manufacturer's value for the drill if the stroboscope is unavailable. Record ω in the Data Table. Convert ω to rad/s and record. Calculate v_s from the equation $v_s = r\omega$ and record.

5. Calculate the Doppler frequencies for approach f_A and recession f_R from source speed v_s found in Step 4 and stationary source frequency found in Step 2; use the following equations:

$$f_A = f\,v\,/\,(v - v_s) \quad \text{(approach)}$$

$$f_R = f\,v\,/\,(v + v_s) \quad \text{(recession)}$$

Record these values and compare then with those found by Step 3.

6. What is "Doppler shift"?

DATA TABLE

Frequency measurments:

Stationary source: SEC/DIV = ____________
DIV/CYCLE = ____________
Rotating Source: SEC/DIV = ____________
DIV/CYCLE = ____________
Determination of v_s: r = ____________
ω = ____________

RESULTS:

Stationary source: T = ____________ s
f = ____________ Hz
Rotating source:

Experimental values:
T_{min} = ____________ s f_A = ____________ Hz

T_{max} = ____________ s f_R = ____________________ Hz

Theoretical values:

ω = ____________ rad/s f_A = ____________________ Hz

v_s = ____________ m/s f_R = ____________________ Hz

SELF CHECK

1. Find the Doppler frequency of an automobile radio tuned to 1500×10^3 Hz (1500 KHz) if (a) the automobile is stationary and the observer approaches at 30 m/s and if (b) the observer is stationary and the automobile approaches at 30 m/s.

2. Find the Doppler frequency heard by a motorist traveling at 20 m/s east, away from the whistle of a locomotive traveling at 75 m/s west if the whistle frequency is 3000 Hz.

3. Explain how beats are produced, and define "beat frequency"; explain how radar speed control works.

4. Find the approach speed of an airplane detected with radar of frequency 3×10^9 Hz if the beat frequency is 990 Hz.

5. Draw a diagram of a shock wave and explain how it is produced; define "Mach cone" and "Mach number."

6. Find the angle included in the bow wave for a toy sailboat moving at 1 m/s if the bow waves have a speed of 25 cm/s.

7. Explain the Doppler broadening of emission and absorption spectral lines.

8. Find observed wavelength of a stellar spectral line of wavelength 6000 A° if the star is receding from Earth at $u = 1.5 \times 10^6$ m/s.

11E1

ALTERNATING VOLTAGE AND CURRENT

When a voltage changes periodically with respect to time, it is referred to as an alternating voltage. This alternating, wave-like voltage, when applied to a complete circuit, will cause an alternating current to flow. Several devices used in electronic circuits react differently to alternating current than they do to direct current. This module compares and contrasts a.c. and d.c. characteristics and describes the effects they produce in various circuit components, such as resistors, capacitors, and coils. In the laboratory, the student will measure peak-to-peak voltage and will compare resistance, inductive reactance, and capactive reactance.

OBJECTIVES

Upon completion of this module, the student should be able to:

- Define the following in one or two sentences:

  Frequency of alternating voltages.
  Peak-to-peak voltages.
  Rms current and voltage.
  Instantaneous voltages.
  Inductive reactances.
  Capacitive reactances.
  Effective voltages.

- List two electrical components that respond differently and one that responds the same to a.c. and d.c.
- Given a capacitor and/or coil, calculate capacitive or inductive reactance. Use Ohm's law to determine value of the current flow at a specific voltage.

406

- Given the appropriate equipment, measure peak-to-peak voltage. Compare resistance, inductance, and capacitive reactance. ance.

DISCUSSION

The use of a graph is often a convenient means of visualizing physical qualities such as velocity, power, rate, etc. Graphical visualization also is useful in visualizing direct and alternating voltages. A graph of a direct voltage versus time is shown in Figure 1, in which a straight line parallel to the time axis indicates that the voltage remains constant over a time period.

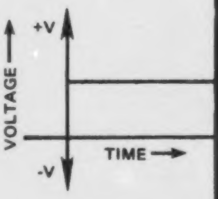


FIG. 1

Conversely, an alternating voltage changes with respect to time, as indicated by the term "alternating." The graph of voltage (Figure 2) follows the form of a trigonometric sine function. Mathematically this graph follows Equation 1.

$$V = V_{max} \sin (2\pi ft) \qquad (1)$$

where

V = voltage at any instant of time t.
V_{max} = voltage at point b in Figure 2.
f = frequency in hertz.
t = time in seconds.

Over a period of time (defined as one period), alternating voltage repeats itself. Figure 2 shows two and one-half cycles of the voltage change. Concept Module 11-0 gives definitions of frequency and period. In Figure 2, the distance from points a to e, from b to f, or, in fact, between any two corresponding points, is the period. In electrical applications, one uses the term "frequency" — the number of times the wave repeats itself each second. The "hertz" unit is used to represent one complete cycle each second, as shown by Equation 2, and thus frequencies are normally given in hertz.

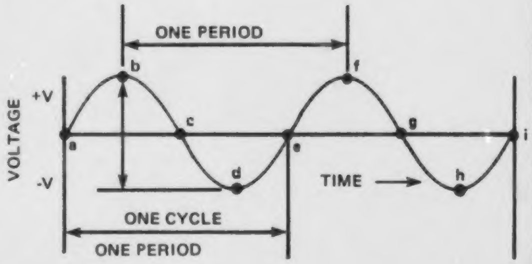


FIG. 2

$$1 \text{ Hz} = 1 \frac{\text{cycle}}{\text{s}} \qquad (2)$$

Commercially-produced electricity in the United States is alternating at a frequency of 60 hertz, although many other frequencies are used in various devices, such as radios, television sets, etc.

EFFECTIVE VOLTAGE

If the value of voltage is always changing, how can one value (such as 120 V) be applied to alternating voltage? The obvious answer to this question would be an "average" value; however, a simple average would be useless, because there are as many negative values of voltage as there are positive values.

A new term, "peak-to-peak voltage," is applied to alternating voltage waves. This very descriptive term defines the voltage measured from the peak of the crest to the bottom of the trough. Peak-to-peak voltage is shown in Figure 2 as the vertical measurement from points b to d. Notice that this value is twice as great as the wave amplitude defined in Module 11-0. As useful as the peak-to-peak value of voltage may be, it does not represent a useful comparison to direct voltage when various effects are measured.

Recall from previous modules using direct current that a direct voltage applied across a given resistance will produce a fixed amount of heat. This heating effect can be used to specify one value for an alternating voltage. A definition therefore could be stated in the following way: one volt a.c. is the alternating voltage that will produce the same amount of heat in a given resistor as one volt d.c. This value is called the "effective voltage." It is the voltage recorded by most voltmeters. Effective voltage is often called rms voltage. The abbreviation rms stands for root mean square, which is a mathematical averaging technique that does not need to be learned here. When the root mean square technique is applied to an alternating voltage sine wave, the rms voltage is found to have the value indicated by Equation 3.

$$V_{rms} = 0.707\ V_{max} \tag{3}$$

where

V_{rms} = effective voltage.

An alternating voltage, therefore, that has an effective voltage (rms) of 1 volt must have a maximum voltage at the crest of 1.414 volts and a peak-to-peak voltage of 2.83 volts. Standard household voltage of 115 effective volts will have a maximum voltage (amplitude) of 162.5 volts and a peak-to-peak value of 325 volts.

The internal heating effect produced in a resistor is the same, therefore, for one volt d.c. as for one rms volt a.c. On the other hand, the inductance coil and the capacitor behave differently when connected to an alternating voltage than when connected to direct voltage. The reaction of these two elements to an alternating voltage can be demonstrated with a simple resistor such as a long, straight, high-resistance wire. When a direct voltage is applied to this wire, a current is produced, constant and proportional to the voltage applied, as described by Ohm's law. An alternating voltage applied to the same wire also will cause a proportional current to flow, but the current will be alternating in step with the voltage. In other words, when the voltage is greatest in one polarity, the current flow will be greatest also; but when the voltage reverses polarity (becomes negative in Figure 2), the current flow will reverse direction as well. This is illustrated graphically in Figure 3. The resistance of the wire is the same regardless of its connection to direct or alternating voltage.

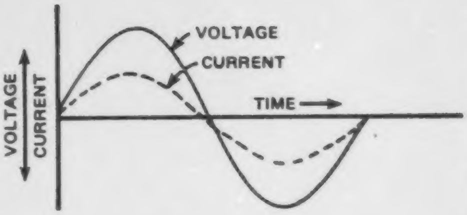


FIG. 3

INDUCTIVE REACTANCE

A coil, on the other hand, reacts against an applied voltage. When voltage is applied to a coil, the current produced establishes a magnetic field around the coil similar to the magnetic field around a bar magnet (Figure 4).

Direct current through the coil establishes a constant and stationary magnetic field. Because the strength of the field depends upon the amount of current through the coil, an alternating current will produce a continually changing magnetic field. As the current alternates, the magnetic lines of force first expand outward, reach a maximum extent, then collapse as the current reverses direction. The continually moving magnetic field near the conducting coil will cause an emf to be induced in the coil according to Lenz's law. The induced voltage is called a counter emf, because it is opposite in polarity to the initially-applied voltage. The counter emf opposes the current flow that produced it. A coil, therefore, opposes the flow of alternating current in a manner different from the way in which it opposes direct current. The extra opposition to alternating current flow is called "inductive reactance." It is measured in ohms and depends upon frequency of the applied voltage and upon construction of the coil (called "inductance"), related in Equation 4.

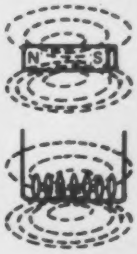


FIG. 4

$$X_L = 2\pi fL \qquad (4)$$

where

X_L = inductive reactances in ohms.
f = frequency of a.c. in hertz.
L = inductance of coil in henries.

This relationship is clarified further by Examples A and B, in which Equation 4 is used to solve problems.

EXAMPLE A: CALCULATION OF INDUCTIVE REACTANCE

Given: A 250-millihenry coil is placed in a 60-hertz a.c. circuit.

Find: The inductive reactance.

Solution: (From Equation 4.)

$$X_L = 2\pi fL$$
$$= 2\pi\ (60\text{ Hz})(250 \times 10^{-3}\text{ henry})$$
$$X_L = 94\text{ ohms}$$

EXAMPLE B: INDUCTIVE REACTANCE IN HIGH-FREQUENCY CIRCUIT

Given: A 250-millihenry coil is placed in a 3000-hertz a.c. circuit.

Find: The inductive reactance.

Solution: (From Equation 4.)

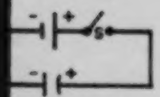
$$X_L = 2\pi fL$$
$$= 2\pi\ (3000\text{ Hz})(250 \times 10^{-3}\text{ henry})$$
$$X_L = 4710\text{ ohms}$$

CAPACITIVE REACTANCE

A capacitor also behaves differently in an alternating-current circuit than in a direct-current circuit. When the switch in the d.c. circuit shown in Figure 5a is closed, the capacitor simply charges to the voltage of the source and then prevents additional current flow, because the capacitor voltage opposes the source voltage. A capacitor, therefore, acts as an infinite resistance (open circuit) to direct current.

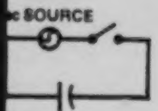
In the a.c. circuit shown in Figure 5b, the capacitor first charges to a maximum, then discharges as the voltage of the source decreases. When the source voltage reverses, the discharging capacitor actually assists the source, which causes the capacitor to offer very little opposition to rapidly-changing current. The opposition of a capacitor, called "capacitive reactance," is given by Equation 5:

$$X_C = \frac{1}{2\pi fC} \qquad (5)$$



dc CAPACITOR CIRCUIT

(a)



ac CAPACITOR CIRCUIT

(b)

FIG. 5

where
X_C = capacitive reactance in ohms.
f = frequency of a.c. in hertz.
C = capacitance of capacitor in farads.

Equation 5 is clarified further by Examples C and D.

Both coils and capacitors produce opposition to alternating current that they do not cause in direct-current circuits. The equations presented in this module indicate that reactance, sometimes called "a.c. resistance," depends upon frequency of the alternating current, as well as upon design of the coil or capacitor.

EXAMPLE C: CALCULATION OF CAPACITANCE

Given: A 0.2-μF capacitor connected to a 60-hertz a.c. circuit.

Find: Capacitive reactance.

Solution: (From Equation 5.)

$$X_C = \frac{1}{2\pi f C}$$

$$= \frac{1}{2\pi\ (60\text{ Hz})(0.2 \times 10^{-6}\text{ farads})}$$

$$X_C = 13{,}300\text{ ohms}$$

EXAMPLE D: CAPACITANCE IN A HIGH-FREQUENCY CIRCUIT

Given: A 0.2-μF capacitor placed in a 10,000-hertz a.c. circuit.

Find: Capacitive reactance.

Solution: (From Equation 5.)

$$X_C = \frac{1}{2\pi f C}$$

$$= \frac{1}{2\pi\ (10 \times 10^{3}\text{ Hz})(0.2 \times 10^{-6}\text{ F})}$$

$$X_C = 80\text{ ohms}$$

SUMMARY

Alternating currents and voltages exhibit a wave-like, or alternating, quality. Frequency is the number of complete cycles per second, (1 Hz - 1 cycle per second). One volt a.c. is the amount of voltage that will produce the same amount of heat in a given resistor as one volt d.c. A peak-to-peak voltage of 2.83 V is required to equal one d.c. volt. The behaviour of d.c. and a.c. voltages differs in circuits having inductive reactance or capacitive reactance.

LABORATORY

MEASURING A.C. VOLTAGE

MATERIALS

Oscilloscope
Function Generator
Digital multimeter or FET multimeter
6 volt dry cell battery (B)
SPST-normally open momentary contact
  switch (S)
150 ohm, 1/2 watt resistor (R)

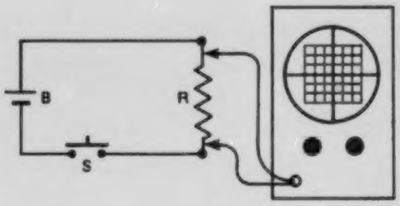


FIG. 6

PROCEDURE

In this experiment an oscilloscope is used. A complete understanding of operation of the oscilloscope is unnecessary at this time. Module 11E3, page 425, gives more information on this instrument. With the aid of the instructor, the student will calibrate an oscilloscope for correct voltage readings.

1. Connect circuit as shown in Figure 6. Set oscilloscope controls: vertical amplifier to 1 volt/division and sweep time to 500 microseconds/division.

2. Draw three sets of xy-axis, labeling the y-axis as voltage and the x-axis as time. Repeatedly open and close switch (S) by pushing and releasing the button. Observe the resulting pattern and draw on the first set of axis.

3. Remove oscilloscope, battery, and switch from circuit and connect the function generator across the resistor. Set function generator: output frequency of 60 Hz, wave-form selector to square-wave, amplitude for 8 full vertical divisions on the oscilloscope. Do not alter amplitude control for the rest of this or the next experiment.

4. Readjust the frequency of the function generator to 1 kHz and draw the resulting pattern on the second set of axes. Change wave-form selector to sine-wave output and draw the resulting pattern on the third set of axes. Measure and record the value of voltage, labeling it a peak-to-peak voltage (Vp-p). Calculate rms voltage by Equation 3.

5. Disconnect oscilloscope; connect multimeter and record rms measurement. Does the calculated value agree with the measured value?

RESISTANCE, INDUCTIVE REACTANCE, AND CAPACITIVE REACTANCE

MATERIALS

Oscilloscope
Function generator (M_1 or M_2)
Digital or FET multimeter - two
6 volt dry cell battery (B)
SPST switch (S)
150 ohm, 1/2 watt resistor (R)
Microfarad capacitor (C)
25 millihenry inductor (L)

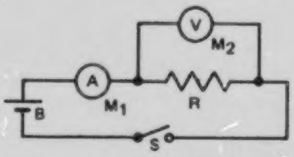


FIG. 7a.

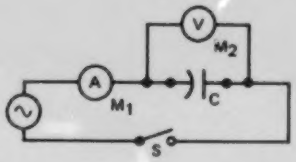


FIG. 7b.

PROCEDURE

1. Connect circuit as shown in Figure 7a. Set multimeter controls: M_1 to d.c. current and M_2 to d.c. voltage.

2. Close switch S; measure and record the voltage and current. From these values calculate R by use of Ohm's Law.

3. Open switch S, connect circuit 7b by replacing the battery with the function generator and the resistor with a capacitor. Set multimeter controls: M_1 to a.c. current and M_2 to a.c. voltage.

4. Set function generator frequency output to 1 kHz sine-wave output. Repeat Step 2 above, only record the calculated value as X_C (capacitive reactance); note frequency.

5. Change output frequency to 10 kHz and repeat Step 2 - recording X_C and frequency.

6. Open switch S. Replace C with inductor L and repeat Steps 4 and 5 above. Record the calculated values as X_L and note frequencies.

7. Using Equations 4 and 5, calculate X_C and X_L. How do these values compare with the values derived from steps 4 through 6 above?

8. Open switch S. Replace M_2 with the oscilloscope*; close S and observe effect on pattern presented on oscilloscope by varying frequency of function generator; record observations.

9. Open switch S. Replace L with C; repeat Step 8.

10. Open switch S. Replace C with R, repeat Step 8.

* Warning - DO NOT connect an oscilloscope and a digital multimeter in parallel to make measurements. Such action can endanger equipment and personnel.

SELF CHECK

1. Define in one or two sentences the following:

    Frequency of alternating voltages.
    Peak-to-peak voltages.
    Rms current and voltage.
    Instantaneous voltages.
    Inductive reactances.
    Capacitive reactances.
    Effective voltages.

2. The peak-to-peak voltage measured with an oscilloscope is 150 volts; what corresponding rms value should be recorded by a voltmeter?

3. List two electrical components that respond differently and one that responds the same to a.c. and d.c.

4. Calculate the capacitive reactance in ohms of a 2-μF capacitor at 5000 Hz.

5. If a 1-μF capacitor is connected to a 117-volt, 60-hertz line, how much current will flow?

6. A 1-henry coil is placed in a 117-volt, 60-hertz line; what is the inductive reactance and how much current will flow?

7. At what frequency will the inductive reactance of a 1 mH inductor be equal to the capacitive reactance of a 1 μF capacitor?

PHASE RELATIONSHIPS IN AC CIRCUITS

PHASE RELATIONS

Direct current encounters opposition from the resistance of conductors. Alternating (wave-like) current encounters the same resistive opposition; but in addition, alternating current experiences opposition in the form of capacitive and inductive reactances. These reactances cause the current and voltage to be out of step with each other. Out-of-step problems cause loss of power in alternating-current circuits.

This module investigates the details of capacitive and inductive reactances, their measurements, and how they affect alternating-current circuits. In the laboratory, the student will determine the impedance, phase angle, and power factor of RC, RL, and RLC circuits.

OBJECTIVES

Upon completion of this module, the student should be able to:

- Draw sine wave sketches, showing the phase relationships between current and voltage in R (purely resistive), RL (resistive-inductive), RC (resistive-capacitive), and RLC (resistive-inductive-capacitive) series circuits.
- Draw and interpret vector diagrams of the phase relationship in RC, RL, and RLC circuits; and determine phase angle and impedance.
- Define these terms:

  Impedance.
  True power.
  Apparent power.

Power factor.
Phase angle.
Resonant circuit.

- Given a series a.c. circuit, calculate the value of additional circuit components necessary to produce resonance (power factor of one).

- Given the appropriate equipment, determine the impedance, phase angle, and power factor from current and voltage measurements taken in RLC, RL, and RC series circuits.

DISCUSSION

PHASE RELATIONS IN RESISTIVE CIRCUITS

In a.c. circuits that contain only resistance, the current flow is in phase with the voltage. The current responds immediately to the voltage, as shown graphically in Figure 1. The power at any instant is the algebraic product of the voltage and current at that instant (Equation 1). Power varies according to the graph in Figure 2.

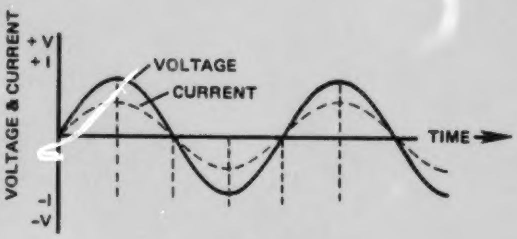


FIG. 1

$$P = VI \qquad (1)$$

where
P = power.
V = voltage.
I = current.

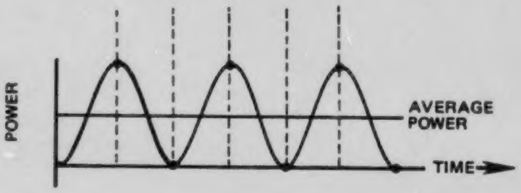


FIG. 2

Notice, also, the value of "average power" given in Figure 2. The average power is the value of the constant power that would produce the same effect as the varying power. Notice that when the current and voltage are in phase, the power always is a positive value.

PHASE RELATIONS IN INDUCTIVE CIRCUITS

When current through a coil varies, as it must in an a.c. circuit, it produces a magnetic field that expands as the current increases and collapses as it decreases. This expanding and collapsing magnetic field, in turn, causes a counter voltage, which opposes change in the flow of current in the coil. The additional opposition to current flow is called "inductive reactance" and is expressed in Equation 2:

$$X_L = 2\pi fL \qquad (2)$$

where
X_L = inductive reactance in ohms.
f = frequency of a.c. in Hz.
L = inductance of coil in henries.

Not only does the coil resist a changing current, but it does so in such a way as to cause the current and voltage to be out of phase; that is, out of step in time. In a purely inductive circuit (one containing only a coil with no resistance), the current will lag behind the voltage by one-fourth a period, or 90° (π/2 radians), as shown graphically in Figure 3. Figure 4, a graph of power versus time, shows the power dissipated by the inductance. The power is positive when both the current and voltage are positive or when they are both negative, but the power is negative when current and voltage are of opposite polarity. The variation of power shown in Figure 4 leads to a zero-average power dissipation by the coil.

FIG. 3

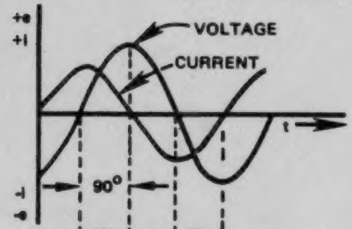
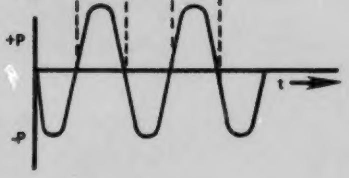


FIG. 4



PHASE RELATIONS IN RL CIRCUITS

A RL circuit is one that contains both resistance and inductance. Since a coil made of wire cannot be free of all resistance, a RL circuit is a more realistic circuit than is the purely inductive circuit represented in Figures 3 and 4.

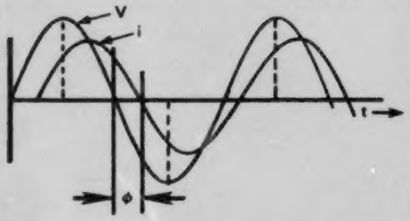


FIG. 5

A circuit containing both resistance and inductance will develop a phase angle difference (ϕ) of between 0° and 90° between voltage and current (Figure 5).

A RL circuit is shown schematically in Figure 6. To determine the total opposition to current, since this is a series circuit, one might expect to add X_L and R. This, however, is incorrect, because of the phase difference between the current and voltage. A procedure identical to that of adding vectors can be implemented. This procedure will give the correct total opposition to alternating current, called "impedance," and the correct phase angle between voltage variation and current variation.

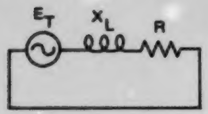


FIG. 6

The procedure is as follows:

(1) On an x-y graph, plot the value of resistance R (in ohms) along the positive x-axis, according to an appropriate scale. This plot will be the horizontal vector R. Plot value of X_L--same scale--along the y-axis.

(2) Use the parallelogram method to add X_L and R. The resultant will be the total opposition (impedance), and the angle between the impedance vector and the resistance vector is the phase angle.

This procedure is demonstrated in Example A.

EXAMPLE A: PHASE ANGLE AND IMPEDANCE OF RL CIRCUIT

Given: A 1500-Hz (hertz) a.c. circuit contains a 30Ω (ohm) resistor and a 2 mH (millihenry) coil.

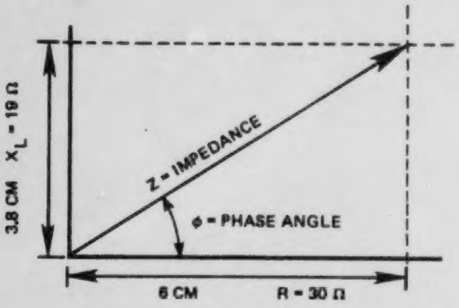
Find: Phase angle and impedance.

Solution:

$$X_L = 2\pi fL$$

$$= 2\pi\ (1500\text{ Hz})(2 \times 10^{-3}\text{ H})$$

$$X_L = 19\Omega$$



SCALE: 1 CM = 5 Ω

From the vector diagram, length of the impedance vector is 7.1 cm; therefore, 7.1 cm x 5Ω/cm = 35.5Ω. Also from vector diagram —

$$\tan\phi = \frac{X_L}{R}$$

$$= \frac{19}{30}$$

$$\tan\phi = 0.633$$

$$\therefore\ \phi = 32.3°$$

Thus, current lags voltage by 32.3°.

POWER IN A RL CIRCUIT

When a voltmeter and ammeter are connected in an a.c. circuit, they record the rms values of voltage and current. If these values are multiplied together, one apparently would find the power dissipated in the circuit. The product of rms voltage and current gives what is called "apparent" power. If the voltage and current were in phase, this product (apparent power) also would be the power dissipated by the circuit. But since the voltage and current are out of phase by an amount equal to the phase angle, the apparent power is greater than the true power. The true power can be determined by Equation 3 in an a.c. circuit when the phase angle is known.

$$P_T = V_{rms} \, I_{rms} \cos \phi \qquad (3)$$

where

P_T = true power.
ϕ = phase angle (power factor).
$V_{rms} \, I_{rms}$ = apparent power.

True power in a RL circuit is illustrated graphically in Figure 7.

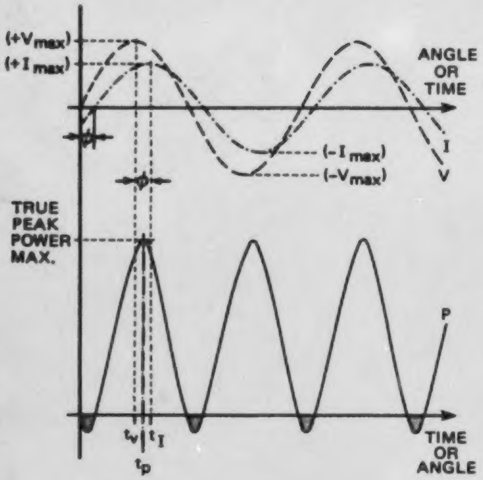


FIG. 7

Notice that when both current and voltage are positive or negative, a positive value for power results; but when either current or voltage is negative, and the other one positive, a negative power results. In Figure 7, more positive power is present than negative power; thus, the average power expended is positive as in a pure resistance. Compare the power graph in Figure 7 with the graphs in Figures 2 and 4. From these graphs and from Equation 3, the true useful power obviously can vary from zero in a pure inductive circuit to a maximum amount in a pure resistive circuit as the phase angle changes from 90° to 0°.

The term "cos ϕ" in Equation 3 is called the "power factor." It relates the true and apparent power in a given circuit. This relation can be seen when Equation 3 is rearranged to become Equation 4.

$$\cos \phi = \frac{P_T}{P_A} \qquad (4)$$

where

P_A = apparent power = $V_{rms} I_{rms}$

True power is often measured in units of watts; apparent power is expressed in units of volt-amps.

PHASE RELATIONS IN CAPACITIVE CIRCUITS

When a capacitor is connected in an a.c. circuit, the opposition presented to the alternating current is called "capacitive reactance," the value of which can be determined from Equation 5.

$$X_C = \frac{1}{2\pi fC} \qquad (5)$$

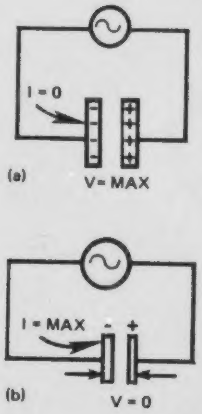


FIG. 8

where

X_C = capacitive reactance in ohms.
f = frequency in hertz.
C - capacitance in farads.

When an alternating voltage is applied to a capacitor, current flows easily to the capacitor plates until enough charge is collected on the plates to build up a significant capacitor voltage. Then, as the capacitor voltage is large, it prevents additional current from flowing to the capacitor plates (Figure 8).

Graphically, the phase relations in a capacitive circuit are illustrated in Figure 9. In this case, the current is said to "lead" the voltage by 90°. (Compare with Figure 4.)

An analysis similar to that for power in inductive circuits can be used for capacitor circuits. The power graph in Figure 10 is the result: the average power dissipated by a capacitor in an a.c. circuit is zero. This relationship follows from Equation 3.

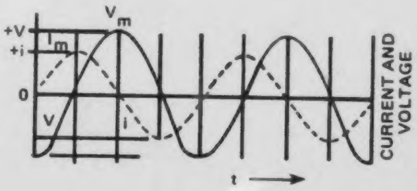


FIG. 9

$$P_T = V_{rms}\ I_{rms} \cos \phi \qquad (3)$$

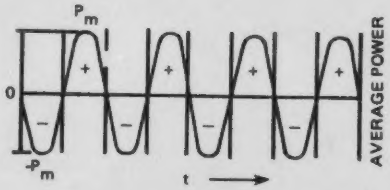


FIG. 10

where

cos ϕ = phase angle (90° in a purely capacitive circuit).
$P_T = V_{rms}\ I_{rms} \cos 90°$.
$P_T = V_{rms}\ I_{rms}\ (0)$.
$P_T = 0$.

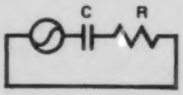
As a consequence of the above values, neither pure inductances nor pure capacitances dissipate any power.

PHASE RELATIONS IN RESISTANCE AND CAPACITANCE CIRCUITS

The phase angle in an a.c. circuit containing both resistance and capacitance will be between 0 and 90°. Vector diagrams can be drawn to determine precisely the impedance and the phase angle; the procedure is similar to that outlined for inductive reactance in Example A. The only modification in the procedure is this: the capacitive reactances must be plotted along the negative y axis to account for the fact that current leads by 90° in a capacitive circuit, but lags by 90° in an inductive circuit. The inductive reactances and capacitive reactances, therefore, are plotted 180° from each other. (Refer to Example B, which consists of an RC circuit problem and solution.)

EXAMPLE B: PHASE ANGLE AND IMPEDANCE OF AN RC CIRCUIT

Given: A series circuit consists of a 1000Ω resistor and a 1 μF (microfarad) capacitor connected to a 117-V (volt) 60-Hz source.



Find: Phase angle and total impedance of circuit.

Solution:

$$X_C = \frac{1}{2\pi fC}$$

$$= \frac{1}{2\pi\ (60)\ (1 \times 10^{-6} f)}$$

$$X_C = 2650\Omega \text{ (approximate)}$$

Verify the result mathematically:

$$1000^2 + 2650^2 = z^2$$

$$z = 2832\Omega$$

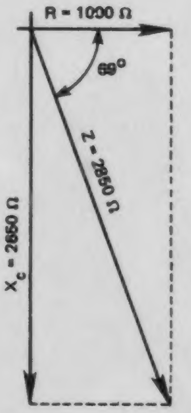
The phase angle is determined:

$$\tan \phi = \frac{X_C}{R}$$

$$= \frac{2650}{1000}$$

$$= 2.65$$

$$\tan \phi = 69°$$



SCALE: 1 CM = 500 Ω

Determine both the true and the apparent power in the circuit:

$$I_{rms} = \frac{V_{rms}}{z}$$

$$I = \frac{117V}{2832\Omega}$$

$$I = 0.04 \text{ A (amps)}$$

$$P_A = V_{rms}\, I_{rms}$$

$$= (117 \text{ V})(0.04 \text{ A})$$

$$P_A = 4.68 \text{ VA}$$

$$P_T = (4.68 \text{ VA})(\cos 69°)$$

$$P_T = 1.68 \text{ watts}$$

PHASE RELATIONS IN SERIES RLC CIRCUITS

Phase relations in a.c. circuits containing capacitive reactances, inductive reactances, and resistance are a simple combination of the concepts presented in the sections on RL, RC, and pure resistance circuits. In fact, the analysis of an RLC circuit will be identical to the analysis of an RL circuit if the inductive reactance exceeds the capacitive reactance. Similarly, an RLC circuit will require the same analysis as a RC circuit if the capacitive reactances are larger than the inductive reactances. If the capacitive and inductive reactances are equal, a special condition called "resonance" exists. A resonant RLC circuit is analyzed in the same way as is a purely resistive circuit, as far as phase relations are concerned.

Consider a series RLC circuit with capacitive reactance of X_C ohms, and inductive reactance of X_L ohms, and a resistance of R ohms. Determine the impedance and phase angle in the circuit.

The following equations are developed for total reactance by use of basic trigonometric and algebraic expressions:

$$X_T = X_L - X_C \quad (6)$$

$$z = \sqrt{R^2 + (X_L - X_C)^2} \quad (7)$$

$$\tan \phi = \frac{X_L - X_C}{R} \quad (8)$$

NOTE: The quantity $X_C - X_L$ simply indicates a difference between the two values. The terms in the equation can be revised, as long as the student analyzes the circuit to determine if it is a capacitive or an inductive circuit.

$$\cos \phi = \frac{R}{z} \qquad (9)$$

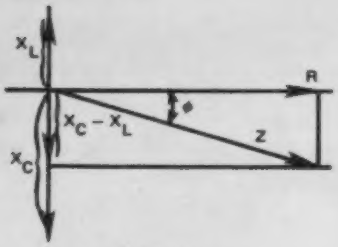


FIG. 11

When X_L and X_C are equal, the value of z will be a maximum equal to R. (See Figure 11.) Since z is minimum at this resonant condition, the current flow will be maximum, and the phase angle equal to zero. At resonance, the relationships are as indicated by Equation 10.

$$X_L = X_C$$

$$2\pi fL = \frac{1}{2\pi fC}$$

$$4\pi^2 f^2 = \frac{1}{LC}$$

$$\text{or} \quad 2\pi f = \frac{1}{\sqrt{LC}} \qquad (10)$$

Besides being point of maximum current flow, a resonant circuit contains the greatest possible power factor, since —

$$\phi = 0$$

$$\cos \phi = 1.$$

When Equation 4 is applied, the apparent power and the true power obviously are equal at resonance. This situation is desirable in commercial power distribution. A power company desires to have the largest possible power factor to provide efficient transmission. Often, a power company will introduce large values of capacitive reactance into a transmission line. This increase compensates for the inductive reactance introduced by large coils in motors or transformers.

Equation 10 also is used in electronic circuits in which the capacitive or inductive reactances are tuned to a desired frequency. This principle is employed in a radio tuning circuit.

SUMMARY

The phase angle ϕ is a measure of how much the voltage leads or lags behind the current in an a.c. circuit. The causes for voltage and current being out of phase stem from the opposition of capacitive and inductive reactances. Loss of power results.

In purely resistive circuits, current is in phase with voltage. Inductive circuits (containing a coil) exhibit "inductive reactance," or opposition to current, which causes the current to lag behind the voltage by 90°. Circuits containing both resistance and inductance exhibit a phase angle of between 0° and 90°, and the "true power" is less than the expected power ("apparent power") of rms voltage multiplied by current. Capacitive reactance causes current to lead voltage by 90°. In resistance/capacitance circuits, the phase angle is between 0° and 90°. Circuits containing resistance, inductance, and capacitance are in "resonance" if the capacitive and inductive reactances are equal. If the inductive reactance exceeds capacitive reactance, circuit analysis is identical to that for RL circuits, but will be identical to that for RC circuits if capacitive reactance exceeds the inductance reactance.

DETERMINATION OF IMPEDANCE, PHASE ANGLE, AND POWER FACTOR

MATERIALS

Resistor, 100 ohms, 1 watt (not wire bound).
Choke coil, 500 mH.
Capacitor, 10 µF, 20 volt rating.
Two switches, SPST.
One a.c. ammeter, 0-0.1 ampere range.
One VTVM.
One low-voltage, 60-Hz power supply, 0-15 volts; or one variable autotransformer.

PROCEDURE

Set up the following circuit illustrated in Figure 12. With switches s_1 and s_2 open, current must go through both the inductor and the capacitor, thus providing a RLC circuit. When s_1 is closed and s_2 open, an RC circuit is produced; with s_1 open and s_2 closed, an RL circuit is available.

Measurements of current and voltage now will be taken and values plotted on a graph. Inductive reactance, capacitive reactance, and resistance will be determined from slope of the graphs; and values of impedance will be compared with the value determined by using Ohm's law.

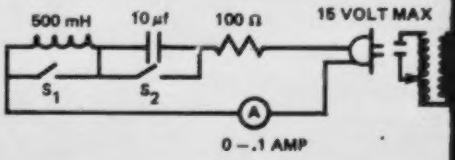


FIG. 12

RLC CIRCUIT

1. With switches s_1 and s_2 open, slowly increase power-supply voltage in uniform increments. Measure voltages across components with VTVM. Record voltage across coil, capacitor, and resistor for each series current reading registered by the ammeter. Complete Data Table 1.

2. Plot voltage versus current for each component on graph in Data Table 1, with current values along horizontal axis and voltage on vertical axis. This plot will produce three lines. Lines may not be perfectly straight because of changes in the components as current increases.

Slope of the graphs will be found from —

$$\text{slope} = \frac{\text{change of voltage}}{\text{change of current}}$$

$$\text{slope} = \frac{\Delta V}{\Delta I}$$

Since $\Delta V/\Delta I$ is the measure of resistance in Ohm's law, the slope of the graph represents opposition to current flow. The following equations can be stated to extend this argument:

$$\frac{\Delta V_L}{\Delta I} = X_L = \text{slope of inductive graph}$$

$$\frac{\Delta V_C}{\Delta I} = X_C = \text{slope of capacitive graph}$$

$$\frac{\Delta V_R}{\Delta I} = R = \text{slope of resistive graph}$$

3. Apply experimental values for X_C, X_L and R in Equation 7 to determine the circuit impedance.
4. Use Equation 8 to determine the phase angle in the circuit.
5. Use Equation 9 to find the power factor of this circuit.

RL CIRCUIT

Close switch s_2 to complete a RL circuit, and perform the same measurements and calculations as in previous procedure on RLC circuit. Use Data Table 2.

RC CIRCUIT

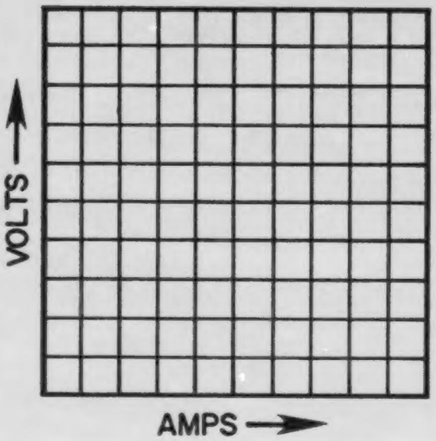
Close switch s_1 and open s_2 to produce an RC circuit, and repeat the same measurements and calculations as in procedure on RLC circuit. Use Date Table 3.

Question for discussion: How can the power factor of an inductive circuit be increased?

DATA TABLE 1: RLC CIRCUIT

RLC CIRCUIT

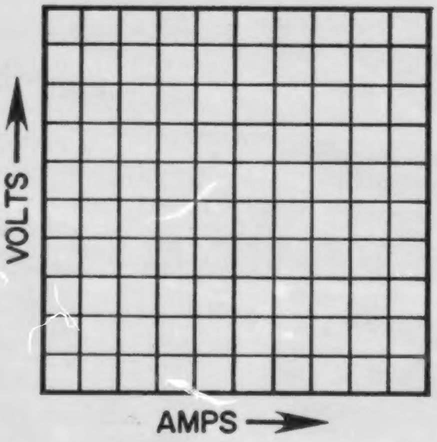
I	V_L	V_C	V_R



DATA TABLE 2: RL CIRCUIT

RL CIRCUIT

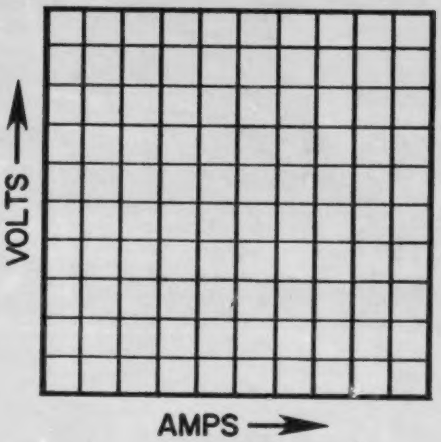
I	V_L	V_R



DATA TABLE 3: RC CIRCUIT

RC CIRCUIT

I	V_C	V_R



SELF CHECK

1. Draw sine wave sketches showing the phase relationship between current and voltage in R, RL, RC, and RLC series circuits. Completely label each sketch, including the phase angle.

2. Define the following terms:

   Impedance.
   True power.
   Apparent power.
   Power factor.
   Phase angle.
   Resonant circuit.

3. Use a vector diagram drawn to scale to determine the impedance and phase angle in a 1500 Hz circuit containing a 300 mH coil and a 700 ohm resistor in series.

4. A 50 millihenry coil is connected to a 60 Hz, 117 V a.c. line containing 80 ohms of resistance. Determine the phase angle, impedance, power factor, true power and apparent power in the circuit.

5. What value of capacitor should be placed in the above circuit to produce a series resonant circuit?

6. What is the phase angle in a circuit if the total capacitance of the circuit is 20 μF, the total inductance is 30 mH, the total resistance is 130 Ω and the circuit is fed by a 60 Hz line?

7. What is the product of the inductance and the capacitance in a circuit if the resonant frequency of the circuit is 60 Hz?

11E3

OSCILLOSCOPE

This module explains the basics of oscilloscope operation. The oscilloscope is one of the most important electrical devices available for measurement of electrical quantities. It is particularly useful in graphic representations of wave forms that may occur at high frequencies. This module also explains basic operation of a dual-trace oscilloscope, its capability, and its applications. It describes use of Lissajous patterns for frequency determination. In the laboratory, the student will analyze oscilloscope sine waves, phase difference, and Lissajous patterns.

OBJECTIVES

Upon completion of this module, the student should be able to:

- Identify which controls of an oscilloscope operate on the x axis and y axis.
- Describe the use of the z axis control.
- Define time per centimeter.
- Discuss the type of oscilloscope required to measure phase difference.
- List circuits affected by the oscilloscope trigger control.
- Given the time-control setting and the time required for one cycle, calculate the frequency of the input signal.
- Calculate the voltage displayed on an oscilloscope screen, given the V/cm control setting and the size of the display.

- Describe the precaution necessary in grounding test equipment.

- Given the appropriate equipment, use a dual-trace oscilloscope to determine the following: (1) peak-to-peak and rms values and frequency of a sinusoidal waveshape; (2) frequency of an unknown sinusoidal waveshape (use Lissajous patterns); and (3) phase differences.

DISCUSSION

CATHODE-RAY TUBE CONSTRUCTION

The cathode-ray tube (CRT) is the major component of a cathode ray oscilloscope (CRO), known variously as an oscilloscope, O'scope, scope, or oscillograph. The CRT is the component around which all other circuitry in the oscilloscope is designed. The CRT itself is an evacuated tube manufactured in various sizes and shapes, depending upon function. Its general construction is shown (Figure 1).

The electron gun assembly contains the cathode, which emits electrons, and several focusing elements that force the emitted electrons into a beam extending from the electron gun assembly forward to the phosphor screen. The electron gun elements both focus the beam and set beam intensity. Electrons in the beam are accelerated further by a high positive potential placed on the phosphor screen. This voltage may vary from 1000 V to 25 kilovolts, depending on size and function of the CRT. When electrons strike the phosphor, they transfer their kinetic energy to the phosphor, which in turn emits photons of energy (visible light) on the face of the tube.

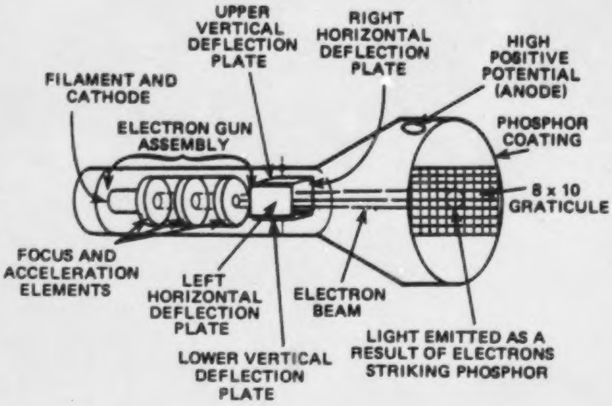

UPPER VERTICAL DEFLECTION PLATE
RIGHT HORIZONTAL DEFLECTION PLATE
HIGH POSITIVE POTENTIAL (ANODE)
FILAMENT AND CATHODE
ELECTRON GUN ASSEMBLY
PHOSPHOR COATING
8 x 10 GRATICULE
FOCUS AND ACCELERATION ELEMENTS
LEFT HORIZONTAL DEFLECTION PLATE
ELECTRON BEAM
LIGHT EMITTED AS A RESULT OF ELECTRONS STRIKING PHOSPHOR
LOWER VERTICAL DEFLECTION PLATE

FIG. 1

In the middle of the CRT are deflection elements. The deflection elements are in pairs, with two horizontal plates causing vertical deflection, and two vertical plates causing horizontal deflection of the electron beam. The plates are charged like opposing plates of a capacitor. A positively charged deflection plate pulls the negative stream of electrons toward it, while a negatively charge deflection plate repels the negative stream of electrons. Each pair of plates receives across it a voltage, which determines either vertical or horizontal deflection of the beam. Thus, the electron beam is deflected as it passes through crossed electric fields. Electrons can also be deflected by magnetic fields. The process described here is called electrostatic deflection.

Figure 2 shows various functional blocks of a single-trace oscilloscope with a triggered time-base for horizontal deflection. Various operator controls are shown beside their corresponding circuitry.

The vertical-attenuator control limits the input signal to

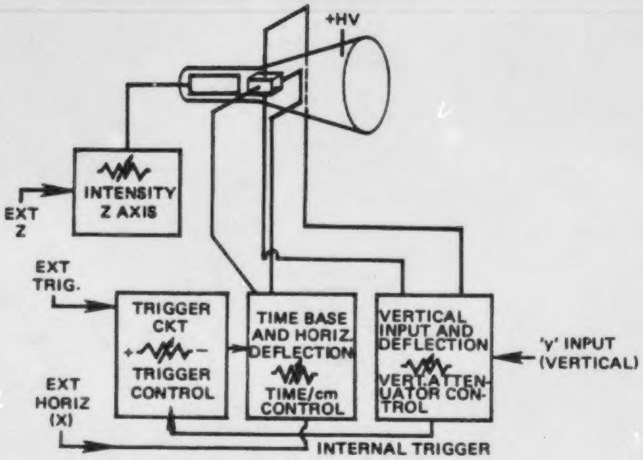

+HV
INTENSITY
Z AXIS
EXT
Z
EXT
TRIG.
TRIGGER
CKT
+ −
TRIGGER
CONTROL
TIME BASE
AND HORIZ.
DEFLECTION
TIME/cm
CONTROL
VERTICAL
INPUT AND
DEFLECTION
VERT.ATTEN-
UATOR CON-
TROL
'y' INPUT
(VERTICAL)
EXT
HORIZ
(X)
INTERNAL TRIGGER

FIG. 2

the vertical deflection circuitry. This determines vertical deflection of the electron beam for a given input voltage. Normally, vertical-attenuator control is provided in steps of 1, 2, and 5. An example of a vertical attenuator control dial is shown (Figure 3). The setting shown (Figure 3) is 50 mV (millivolts)/cm. This value means that for every 50 mV of input signal (the "y" input), the beam will deflect vertically 1 cm on the phosphor face of the tube.

The horizontal control dial for the time base is essentially the same as the vertical attenuator, except that it controls the x axis and is normally in units of time (Figure 4). The setting shown (Figure 4) is 0.1 ms (millisecond)/cm, which means the trace will cover one centimeter in a horizontal direction in 0.1 ms. If the graticule contains 10 cm (centimeters) horizontally, the trace will cover the complete screen in (10 x 0.1) or 1 ms. The x/y input position allows the operator the choice of inserting his own x-axis input. In different applications, this voltage might represent a time or a frequency.

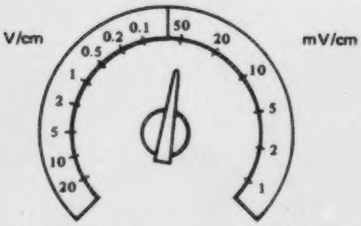

V/cm
mV/cm
0.1
0.2
0.5
1
2
5
10
20
50
20
10
5
2
1

FIG. 3

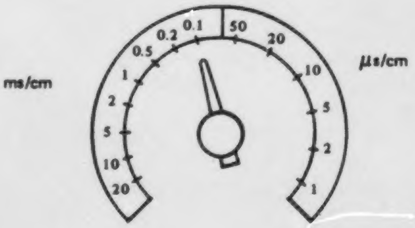

ms/cm
µs/cm
0.1
0.2
0.5
1
2
5
10
20
50
20
10
5
2
1

FIG. 4

The trigger input circuit determines starting point of the trace on the "y" input signal. It can also determine the level of "y" input signal necessary to cause a trace to be triggered. Most oscilloscopes have an automatic trace feature that will allow a trace to appear in the absence of any input circuit signal to the "y" deflection circuitry. Normally, the vertical input circuitry signal ("y") is used for triggering the trace; but most scopes make an external input available to the trigger circuit, allowing an external event to cause triggering, thereby synchronizing the trace to the external event.

The z axis is the intensity circuit that can vary the brightness of the trace. The z axis is normally available with an external input that allows a signal to enhance (or blank) any portion of the displayed signal occuring during the z axis signal.

Listed controls are those found on almost all laboratory oscilloscopes. There are many other control functions available, and they differ from manufacturer to manufacturer and from model to model. A thorough reading of the operation manual for the oscilloscope used is recommended for successful operation.

SINUSOIDAL WAVE ANALYSIS

The pictorial display presented on the oscilloscope by a wave pattern allows measurements of its amplitude and frequency. In addition, characteristics such as linearity, symmetry, and other parameters of the signal are available for diagnosis. Figure 5 represents a typical sinewave (sinusoidal) display on an oscilloscope. The settings of the oscilloscope controls are listed beside the display. These settings are used to determine amplitude and frequency of the wave as follows.

The vertical attenuator setting is 2 V/cm and the horizontal setting is 5 ms/cm. Since peak-to-peak voltage can be observed to span 5 cm, the peak-to-peak voltage is 10 V. The peak voltage is 5 V (2.5 cm) and rms (effective) voltage can be obtained by multiplying peak voltage by 0.707, or approximately 3.5 volts.

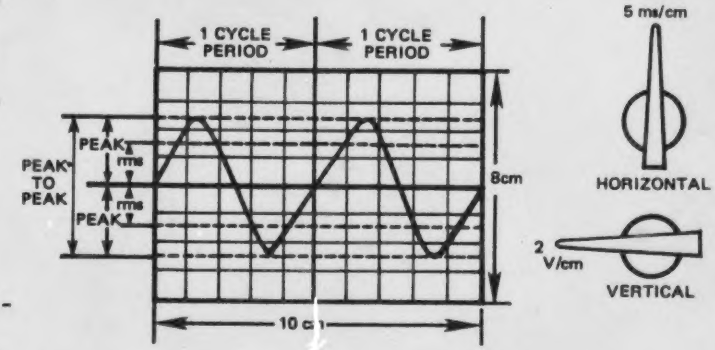


FIG. 5

Frequency of the input signal is determined by counting the number of centimeters spanned by a full cycle of input signal. In Figure 5, a full cycle (one point on the signal to the next corresponding point on signal) takes up 5 centimeters. Thus, the period of the wave is 5 cm x 5 ms/cm or 25 ms. Frequency is determined by taking the reciprocal of the period, which results in 1/0.025 = 40 Hz. Thus, the signal frequency is 40 Hz.

The real value of the oscilloscope becomes apparent when signals other than those of the simple sinusoidal form are displayed. Figure 6 shows some complex waves that can be analyzed from the scope display. Even for such complex wave forms, the oscilloscope presents a display that gives the following information about the signal(s): (1) whether or not the signal is periodic (repeats); (2) signal amplitude; and (3) signal repetition rate (frequency). If several signals are admitted together, the scope displays the resultant, or sum, of the individual signals immediately.

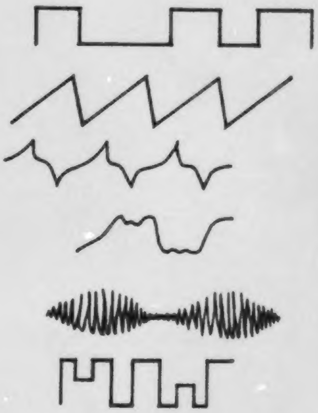


FIG. 6

Such resultants might otherwise require a great deal of mathematics to determine. All the wave forms shown in Figure 6 can be considered as a sum of sinusoidal waves of different amplitudes and frequencies.

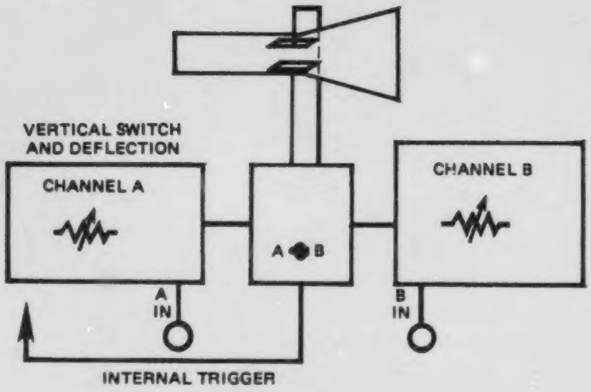


FIG. 7

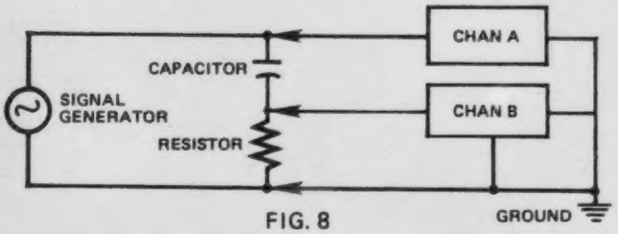


FIG. 8

A dual-trace scope can be used to determine the phase relationship of one signal to another. A dual-trace scope has the same general block diagram as the basic oscilloscope, except that another vertical channel is provided; and a switching circuit is able to display alternately one channel and then the other, giving the illusion of two signals simultaneously displayed. A simplified block diagram is shown in Figure 7. In the circuit shown schematically in Figure 8, the phase difference between the applied voltage and the current in an RC (resistance-capacitance) circuit can be observed. The oscilloscope is connected in such a manner that voltage across the capacitor-resistor combination goes to channel A, and voltage across the resistor (in phase with the current) goes to channel B. Since voltages in an a.c. circuit generally are not in phase, the dual trace oscilloscope can be used to determine phase differences by permitting measurement of the relative shift of the sine waves.

If, for example, resistance R of the resistor (Figure 8) equals capacitive reactance X_C of the capacitor, there will be a 45° phase shift between applied voltage and voltage across a pure resistance. Note: Caution must be observed, in that most oscilloscopes have a common, or chassis, ground (as do most signal generators). When connected in a signal circuit, care must be taken to ensure that all ground points correspond to the ground reference point of the circuit.

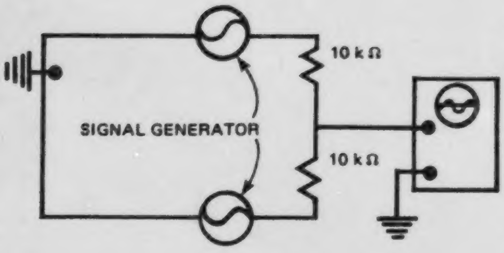


FIG. 9

When a single trace with two mixed sources is used, an algebraic result can be represented. Figure 9 shows the mixing of two frequency sources. Patterns of reinforcement, cancellation, and harmonic interference can be observed with this method. Reinforcement occurs when both signals are in phase. Cancellation occurs when both signals are out of phase.

LISSAJOUS PATTERNS

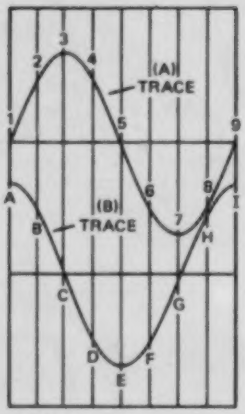


FIG. 10

Lissajous patterns are a method of determining frequency of a signal in comparison to a known one, and to some extent, the phase of a signal in comparison to an unknown signal of the same frequency.

Figure 10 shows two signals of the same amplitude and frequency, but 90° out of phase. When one signal (A) is applied to the y-deflection plate and the other (B) to the x-deflection plate (x-y position of time/cm switch or external horizontal input), the display shown in Figure 11 is obtained. Note that x and y deflections in the circular pattern can be related at each point to deflections shown in the A and B traces. One revolution of the Lissajous pattern corresponds to one period of the waves. If the same frequency signals are in phase (Figure 12a) they give a Lissajous pattern (Figure 12b). If identical frequency signals are 180° out of phase (Figure 13a), they result in a Lissajous pattern (Figure 13b).

More interesting patterns result if the two perpendicular signals are of different frequencies, especially when their ratio is a whole number. Figures 14a and 14b show two sets of such signals, the known signal going to the horizontal plates and the unknown signal going to the vertical plates.

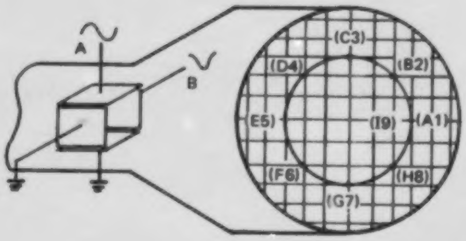
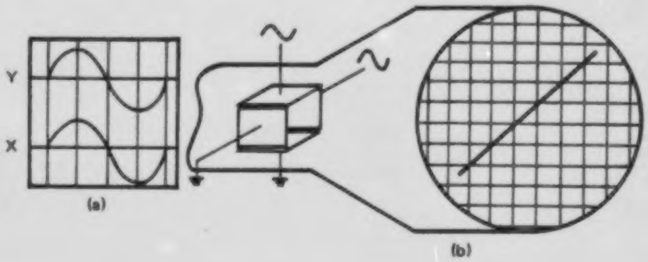
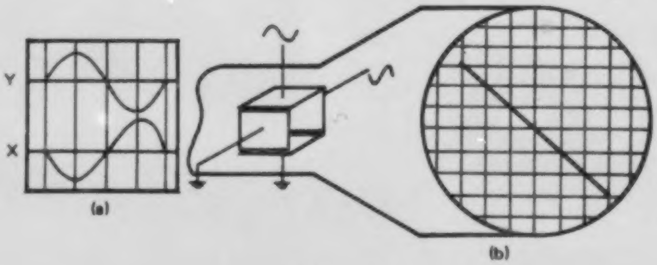


FIG. 11



(a) (b)

FIG. 12



(a) (b)

FIG. 13

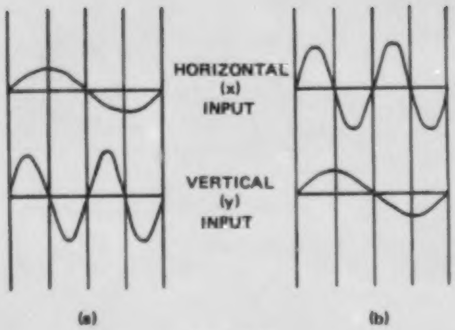

HORIZONTAL
(x)
INPUT
VERTICAL
(y)
INPUT
(a)
(b)

FIG. 14

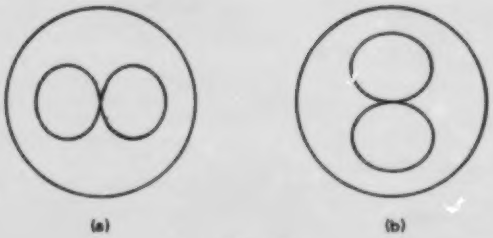


FIG. 15

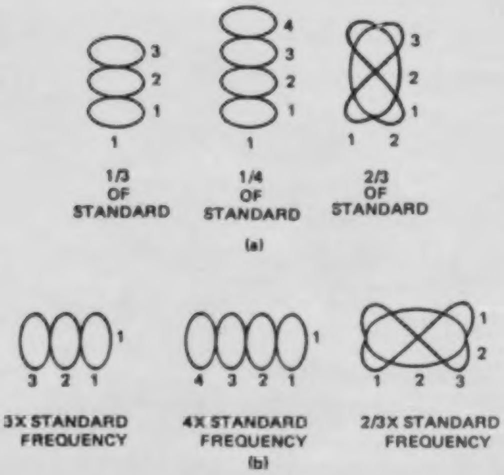

3
2
1
1
1/3
OF
STANDARD
4
3
2
1
1
1/4
OF
STANDARD
3
2
1
1 2
2/3
OF
STANDARD
(a)
1
3 2 1
3X STANDARD
FREQUENCY
1
4 3 2 1
4X STANDARD
FREQUENCY
1
2
1 2 3
2/3X STANDARD
FREQUENCY
(b)

FIG. 16

Figures 15a and 15b show the resultant Lissajous patterns. In one case, the unknown signal has twice the frequency of the horizontal (a); and in the other case, the unknown signal is one-half the horizontal frequency (b). Such patterns are particularly useful in determining frequencies out of the range of the signal generator.

There are many Lissajous patterns; Figures 16a and 16b show several of them, plus methodology for determining input frequency from the pattern. In Figure 16a, unknown is submultiple of known (x) input. The unknown is, therefore, a lower frequency than known input. Count number of loops vertically on one side and divide into standard frequency. Count number of loops horizontally (at top or bottom) and multiply by standard frequency. In Figure 16b, unknown is multiple (higher frequency) of known (x) input. Count number of loops horizontally and multiply by standard frequency. Count number of loops vertically on one side and divide into standard frequency.

SUMMARY

The oscilloscope is a device that utilizes a cathode-ray tube to produce a visual display of the waveforms of electromagnetic phenomena. It is one of the most vital and convenient instruments for the measurement of electrical quantities. The extreme frequency range is one of its principal features; that is, oscilloscopes are capable of displaying voltages having frequencies of hundreds of megahertz. Photographs often are taken of the data displayed.

SINEWAVE ANALYSIS

MATERIALS

Dual trace oscilloscope with a triggered time base and calibrated horizontal sweep with a minimum of 1 μs/cm (microsecond) centimeter) sweep time.

Function generator with variable output amplitude (to exceed 1 V_{rms}) minimum frequency range of 30 Hz to 300 kHz.

PROCEDURE

Set up the circuit shown in Figure 17. Set the controls of the oscilloscope to 0.1 ms/cm on the horizontal time base and 0.1 V/cm on the vertical amplifier. Adjust the trigger level for a steady display, and be sure the input selector is for a.c. signals with the trace centered on the display. Set the output frequency of the function generator to 1 kilohertz with a sine wave output. Now, adjust the output amplitude of the function generator for a display that fills the display 4 cm above and below the center graticle line.

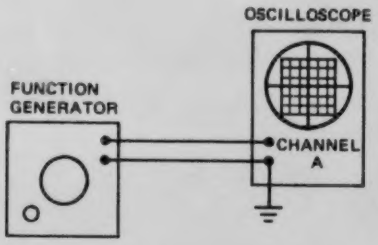


FIG. 17

1. Record the peak voltage as displayed by the oscilloscope.

2. Record the peak-to-peak voltage.

3. Record the rms voltage.

4. Count the number of centimeters for one complete display cycle and multiply this number by the 0.1 ms/cm time base setting. This is one period of oscillation. Calculate the frequency displayed by use of equation

$$f = 1/\text{period} \quad (1)$$

What is the difference, if any, between the calculated frequency and the output setting of the function generator.

PHASE DIFFERENCE

MATERIALS

Same as Procedure 1, and further including:

1-0.01 μF capacitor (molded mylar), 100 V d.c.
1-2 kΩ (2000 ohm) variable potentiometer.

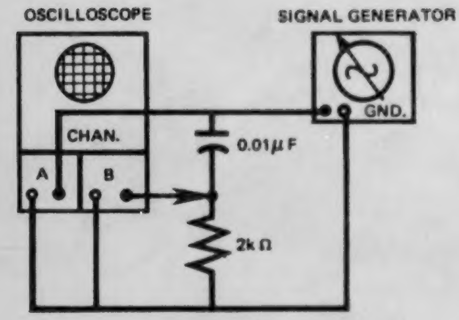


FIG. 18

PROCEDURE

Assemble circuit (Figure 18).

1. Set variable resistor to midrange.

2. Set function generator for 10 kHz; use dual trace, and display on channel A the voltage across capacitor-resistor combination. Triggering should be taken off channel A. Channel B should display voltage across resistor.

3. Set vertical deflection on channel A for 0.5 V/cm, and adjust function generator output to display four centimeters of amplitude.

4. Adjust vertical deflection on Channel B for a viewable signal, preferably four centimeters in amplitude; and display channel B signal below channel A signal on screen.

5. Slightly vary potentiometer. Notice that the phase between the signal on A and signal on B will change. Extreme rotation of control to low resistance will decrease viewable signal beyond usefulness. Use only midrange (1000 ohms) to maximum resistance (2000 ohms). When $X_c = R$ (where X_c = capacitive reactance and R = resitance), there should be a 45° phase difference between both signals.

LISSAJOUS PATTERNS

MATERIALS

Same as Procedure 1, except that two signal generators are required.

PROCEDURE

Assemble equipment (Figure 19).

1. Set both signal generators to approximately the same output signal amplitute and frequency.

2. Set time/cm switch to x-y position (or use external horizontal input on some scopes).

3. Adjust "y" input attenuator until a roughly square display is obtained.

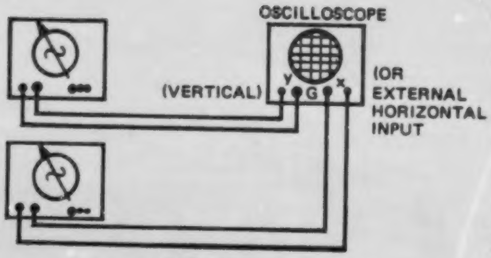


FIG. 19

4. Slightly vary frequency control of one signal generator until a circular trace occurs. This trace means that the two generators are producing the same frequency, but are 90° out of phase. In all probability, the circle will vary from a 45° line running up from left to right (180° out of phase) through 45° running up from right to left (in phase).

5. Adjust signal generator to a multiple of "x" input (horizontal) generator. When multiple is reached, notice pattern reproduced.

6. Try to draw all patterns shown in Figure 16.

SELF CHECK

1. Vertical control operates on (x)(y) axis?

2. Horizontal control operates on the (x)(y) axis?

3. The z axis control affects the ______ of the display.

443

4. Time per centimeter refers to which of the below items?

   Entire horizontal axis.
   Entire vertical axis.
   One vertical division.
   One horizontal division.

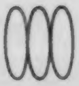
5. A __________ type oscilloscope is required to measure phase difference.

6. The trigger control affects the ___________ circuit.

7. With time/cm control set to 0.2 milliseconds per cm, a waveform takes 7.5 cm to complete one cycle. What is its frequency?

8. With V/cm control set to 0.2 V/cm and a display that fills 6.7 cm, what is the voltage displayed?

9. What precaution concerning grounding of test equipment is necessary in making a proper measurement?



10. If an oscilloscope has an input of 60 Hz to its horizontal plates, what is the input to its vertical plates if the oscilloscope displays the Lissajous pattern shown at the left?

11. If the peak-to-peak voltage displayed on an oscilloscope is 12.7 volts, what is the rms voltage?

12. If 10 cm is the maximum peak-to-peak height that will fit on an oscilloscope's display, what volts/cm setting should be used so that a 2 V volt (peak-to-peak) signal fills the entire display?

11E4

RECTIFIERS

The conversion of alternating current into direct current is necessary in many electronic circuits. A rectifier (diode) accomplishes this conversion. In this module, the principle of rectification will be discussed, and operation of a rectifier in an electrical circuit will be examined. Rectification in fluid and mechanical systems also will be discussed. In the laboratory, the student will assemble circuits containing rectifiers, and will examine their performance with an oscilloscope.

OBJECTIVES

Upon completion of this module, the student should be able to:

- Define, in two or three sentences, an electronic rectifier.
- Draw and label graphs as functions of time for the following:

  Sine wave signal applied to a rectifier and resistor in series.
  Voltage across rectifier.
  Voltage across resistor.
  Current through circuit.

- Describe a bridge rectifier circuit (include a diagram with identification labeling of major items such as a.c. signal source and load resistor).
- Describe a full-wave rectifier circuit for use with a center-tapped transformer (include a diagram with identification labeling or major items such as transformer and load resistor).

- Give an example of rectification in:

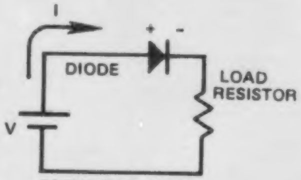
    A fluid system.
    A mechanical system.

- Given the appropriate equipment, construct half-wave and full-wave rectifiers, and measure time-varying voltages on each component with an oscilloscope.

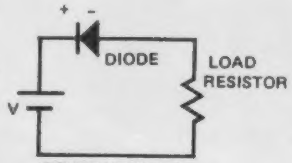
DISCUSSION

ELECTRONIC RECTIFIERS

In electronic terms, a rectifier (also called a diode) is a device that allows a current to flow in one direction only. Figure 1 shows two circuits, one containing a forward- and the other a reverse-biased rectifier. The electronic symbol for a rectifier is an arrow with a vertical bar at the point (→|). The bias of a rectifier is the direction of potential difference across the rectifier. When the back of the arrow is made positive and the point is made negative, the rectifier is forward biased (+ →| -) and represents a low resistance to current flow in the forward direction. A small voltage (approximately 2.3 volts for a solid state silicon rectifier) appears across the diode. The remainder of the applied voltage appears across the resistor.



FORWARD BIASED: CURRENT FLOWS



REVERSE BIASED: NO CURRENT

FIG. 1

When the diode is reverse biased, the back of the arrow is negative and the point is positive (- →| +). In this case, the diode presents an almost infinite resistance and effectively blocks current flow in the circuit. In this direction, most of the voltage appears across the diode. As the resistance is not actually infinite, a small current does flow in the circuit. This is called the leakage current. It is so small, however, that it usually can be ignored.

In the remainder of this module, the forward resistance of diodes will be treated as zero; and the reverse resistance, infinite. This treatment will give a voltage of zero on the diode when it is forward biased and a voltage equal to that of the source when the diode is reverse biased. This treatment of an ideal diode introduces little error in most circuits.

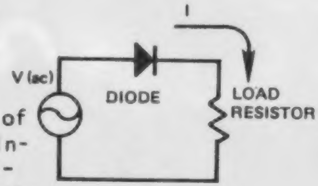


FIG. 2

Figure 2 shows a typical circuit containing a.c. input signal, diode and load. Figure 3 shows the time-varying voltages and currents in diode circuit shown in Figure 2. The source voltage is a sine wave that results in an alternating forward and reverse bias on the diode. When the upper terminal of the a.c. source is positive, the diode is forward biased and has zero resistance. The total voltage will appear across the resistor and will be, at any instant, the product of current in the circuit and resistance of the load resistor. The current direction will be that shown in Figure 2.

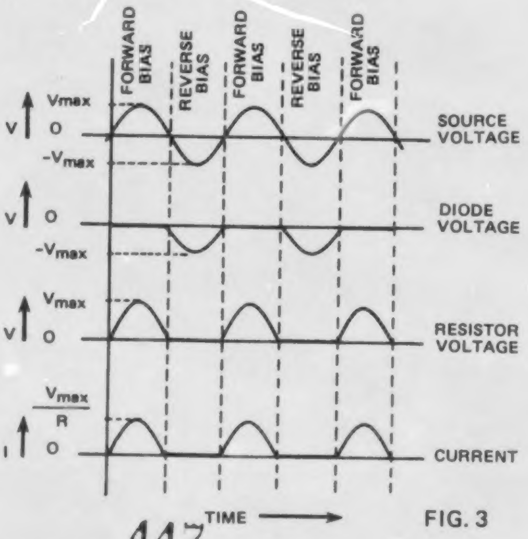


FIG. 3

When polarity of the voltage source (V_{ac}) reverses, the diode becomes reverse biased and provides an infinite resistance. The voltage of the source now appears across the diode. The resistor voltage is zero and the current in the circuit is zero. The variation of current with voltage is shown in Figure 3.

The circuit discussed above is called a half-wave rectifier circuit because the applied signal voltage appears on the load resistor only half the time - during the forward-bias portions of the cycle. This voltage is always positive, although not constant in magnitude, and, therefore, is "rectified" in the form of a pulsating d.c. voltage.

In many applications, the transmission of all applied voltage to the load, and the rectification of it at the same time, is desirable. This procedure is accomplished with the full-wave bridge rectifier circuit incorporating four diodes (Figure 4.)

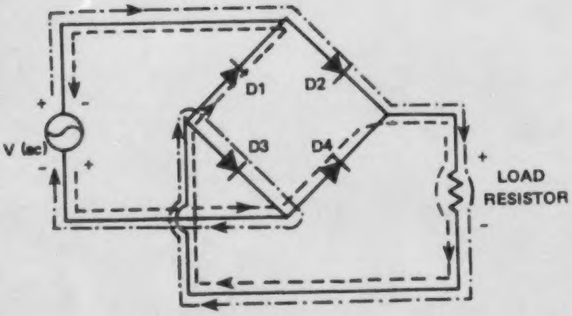


FIG. 4

During the half cycle when the upper terminal of the source is positive, D2 and D3 are forward biased and D1 and D4 are reverse biased. Current flows from the upper terminal of the source through D2, through the load resistor, through D3 and to the lower terminal of the source, as indicated by the dotted line in Figure 4. During this half cycle, D1 and D4 block current flow. During the following half cycle when the lower terminal of the source becomes positive, the bias direction changes for each diode. Current then flows from the lower (positive) terminal of the source through D4, through the resistor, through D1 and to the upper terminal of the source (dashed line). Direction of current flow through the resistor is the same in both half cycles. Figure 5 shows the resulting voltage on the load resistor in Figure 4. The rectified voltage is again a pulsating d.c. voltage, this time without missing half cycles.

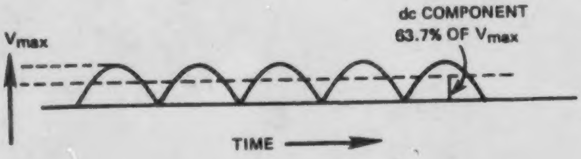


FIG. 5

Figure 6 shows another kind of full-wave rectifier circuit, used with a center-tapped transformer and only two diodes. When point A is positive, D1 is forward biased and current flows as indicated by the dotted line. Because point C is negative, D2 is reverse biased and no current flows through it. On the next half cycle, point A becomes negative and point C becomes positive. D2 is then forward

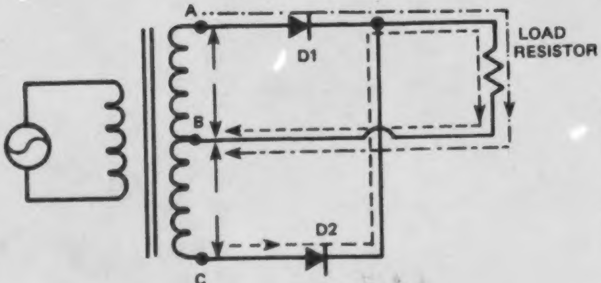


FIG. 6

biased, allowing current to flow as shown by the dashed line. The result is again the full-wave rectified voltage across the resistor shown in Figure 5. This is not a pure d.c. voltage. It can be thought of as containing both d.c. and a.c. components, although the result is never negative. For full-wave rectification, the d.c. component is 63.7% of the maximum voltage. For half-wave rectification, the d.c. component is half this, or 31.8% of V_{max}.

Rectifiers are used in electronic circuits when a.c. currents and voltages must be converted to d.c. This need arises most commonly in power supplies for electronic equipment, in logic and switching circuits, and in demodulation circuits for communication. The principle of rectification is also applied in other energy systems.

MECHANICAL RECTIFIERS

Figure 7 shows a check valve (rectifier) for use in a fluid system. When fluid pressure is applied from the spring end of the valve, both spring and pressure force the ball into the seat; and no flow is present. When pressure is applied from the ball end of the valve, the spring is compressed and fluid passes around the ball. The valve then behaves like a fluid-system diode, allowing flow in one direction only.

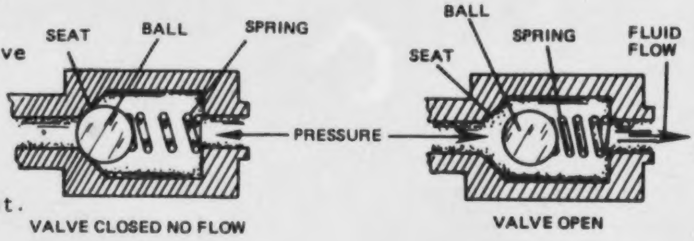


FIG. 7

Figure 8 shows operation of a ratchet (rectifier) in a mechanical system. The hinge in the lever is designed to bend upward only, and the ratchet has a stop that will allow it to swing upward only. When a downward force is applied to the lever, the rack moves upward. The ratchet swings upward to allow one tooth of the rack to pass. When the free end of the lever is moved upward, the ratchet prevents a downward motion of the rack. The hinged end of the lever moves downward past the next tooth of the rack, and the entire process then may be repeated. This rack and ratchet system then operates as a mechanical system diode, allowing mechanical motion in one direction only.

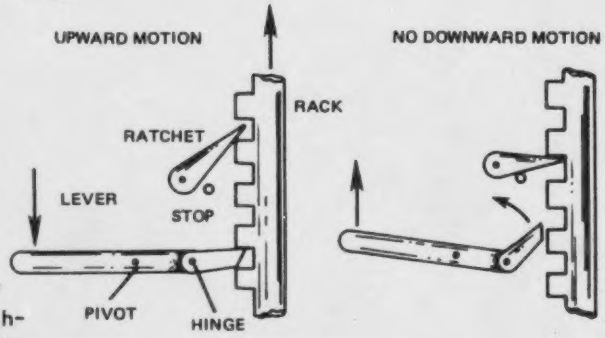


FIG. 8

SUMMARY

Rectifiers allow motion in only one direction. An electronic rectifier (diode) allows current to flow in one direction only, thereby converting alternating current into direct current. Mechanical rectifiers include the check valve (which permits fluid to flow in but one direction) and the ratchet (which limits mechanical motion to one direction).

PERFORMANCE OF RECTIFIER CIRCUITS

MATERIALS

a.c. power cord with clips attached.
Four diodes, 1 A, 100 PIV minimum.
Resistor, 2 W, 1000 Ohms minimum.
Transformer, 24 V, 0.25 A minimum.
Clip leads.
Dual-trace oscilloscope.

Note: A single-trace oscilloscope may be used, but it will not show the phase relationship of rectifier and resistor voltages. These two signals must be viewed separately if a single-trace scope is used.

PROCEDURE

CAUTION: Dangerous voltages are present when this circuit is energized. Do not touch any components while power cord is plugged in. Disconnect power at end of each step.

1. Connect circuit as shown in Figure 9 (do not plug in power cord); set oscilloscope sensitivity for 10 V/cm; set time base for 5 ms/cm. Turn on oscilloscope and obtain a stable trace.
2. Plug in power cord; observe signal on oscilloscope; sketch signal in Data Table. Unplug power cord.
3. Connect circuit as shown in Figure 10.
4. Plug in power cord; observe traces on scope. Trace A is diode voltage; trace B is resistor voltage. Sketch both signals in Data Table. Unplug power cord.
5. Connect circuit shown in Figure 11.
6. Plug in power cord; observe voltage on resistor and sketch in Data Table. Unplug power cord.

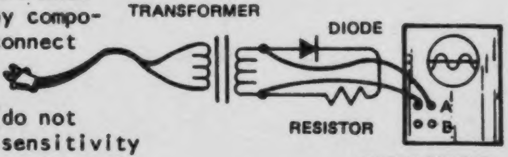


FIG. 9

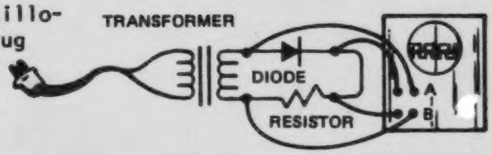


FIG. 10

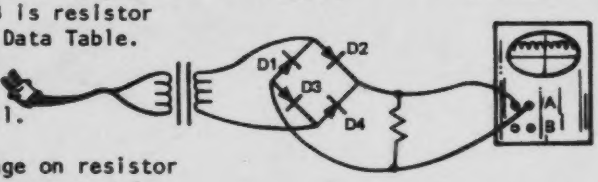
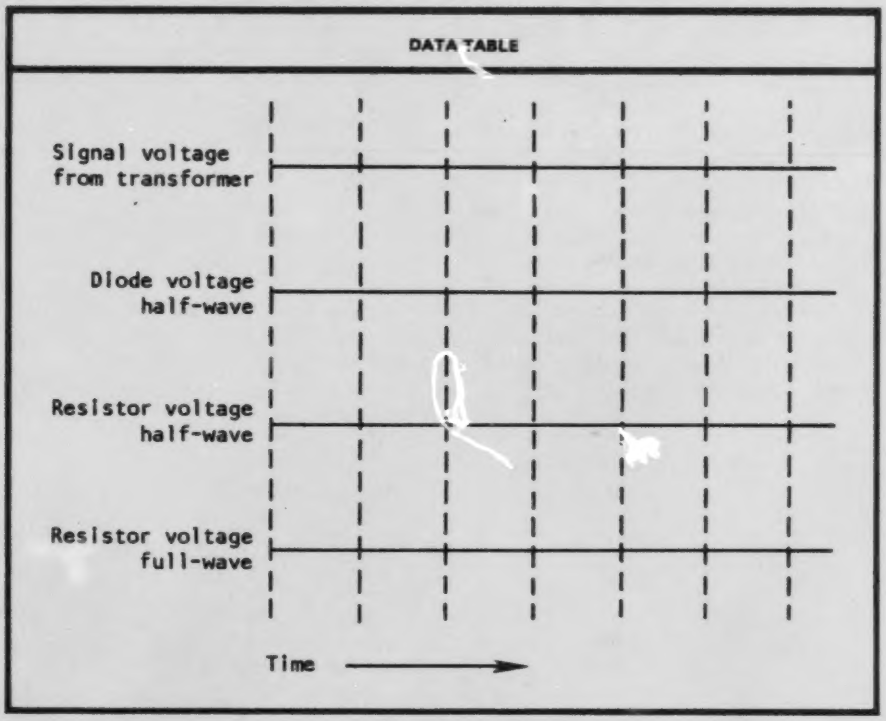
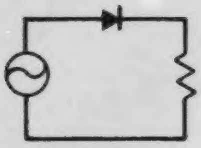


FIG. 11

7. What do the oscilloscope traces for each circuit indicate?



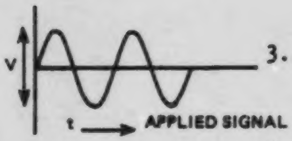
SELF CHECK



1. Define an electronic rectifier.

2. For the circuit at left, draw and label graphs as functions of time for:

   Voltage across rectifier.
   Voltage across resistor.
   Current through circuit.



3. Draw and label a diagram of a bridge rectifier circuit. Show current flow through the circuit for both the positive and negative half-cycle of the signal source.

4. Draw and label a full-wave rectifier circuit for use with

a center-tapped transformer. Show current flow through the circuit for both the positive and negative half-cycle of the signal source.

5. Draw and label diagrams illustrating an example of rectification in:

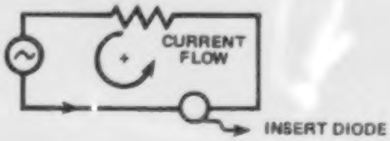
   A fluid system.
   A mechanical system.

6. What is the potential drop across a reverse-biased diode when the applied voltage is 6 V?

7. There is a small potential drop across any real forward-biased diode. If the potential drop across the diode is 2.3 volts, what is the resistance of the diode when a current of 20 mA. flows through the diode?

8. If the voltage from an a.c. source is given as $V = 6 \sin 377\, t$, how much time elapses between the time a diode in the circuit becomes reverse-biased until the time it becomes forward-biased again?

9. In the circuit shown below, indicate which direction the diode should be inserted into the circuit if the diode is to be forward-biased when positive current flows counterclockwise in the circuit.



10. How much does the small resistance (40 Ω) of a forward-biased diode reduce the current in a simple circuit consisting of a 12 volt d.c. voltage source and a load resistance if, before the diode is inserted, the current flowing in the circuit is 8 mA?

TIME CONSTANTS

APPLICATION MODULES

		Page
12F1	Rate of Emptying Tanks	447
12E3*	RLC Circuits	461
12T1	Response Time of Thermocouples	471
12T2	Time Constants of Heating and Cooling	486

* Identifies concept support modules that introduce new physics.

12F1

RATE OF EMPTYING TANKS

Large cylindrical tanks are used widely for storage of water and other fluids in industries that handle water, waste water, chemical processing, and fire control. When these large tanks are drained by gravity, the emptying process resembles the discharge of an electrical series resistance-capacitance combination and, in a similar manner, can be described by a time constant. This module describes the process of emptying a cylindrical tank and discusses the various parameters associated with the fluid flow from the tank. In the laboratory, the student will empty a large cylindrical container to demonstrate the fluid-system behavior described in the module.

OBJECTIVES

Upon completion of this module, the student should be able to:

- Define "fluid capacitance."
- Given a fluid capacitance of a cylindrical tank and a fluid resistance of an outlet, calculate the time constant of the fluid system.
- Given the height of an initial fluid level above an outlet of a cylindrical tank and the time constant of the system, calculate the fluid level at some given time after the tank begins to empty.
- Given the time constant of an emptying cylindrical tank, calculate the approximate time required to empty the tank.
- Given the height of an initial fluid level above an outlet of a cylindrical tank, the elapsed time since the beginning of the process, and the fluid level height after the elapsed time, calculate the time constant of the tank.

- Given the appropriate equipment, determine the rate of emptying tanks.

DISCUSSION

City water supplies usually employ large storage tanks in high places. Water is pumped into the tanks and later flows to consumers in the area. Tanks store the potential energy of the water and are, in a sense, hydraulic capacitors. (Recall: capacitors store electrical potential energy.) Tanks in need of repair must be drained before work crews go inside. Since the mass of the fluid in the tank cannot be emptied instantaneously, but requires some period of time, is is useful to know the time for a transfer of fluid to take place and how the volume of fluid changes with time as the tank empties.

Figure 1 illustrates a large tank, vented to the atmosphere at the top and provided with an outlet near the bottom. The release valve behaves as a variable hydraulic resistor, since it will control the rate at which the tank will empty. (Recall: electrical resistors control the rate of charge flow.) When the valve is fully open, the tank will empty at its fastest rate. When the release valve is closed partially, the result is a larger resistance to the flow, which slows down the rate at which the tank empties. The tank itself is a fluid capacitor with stored-up (potential) fluid energy directly proportional to the volume of fluid above the outlet.

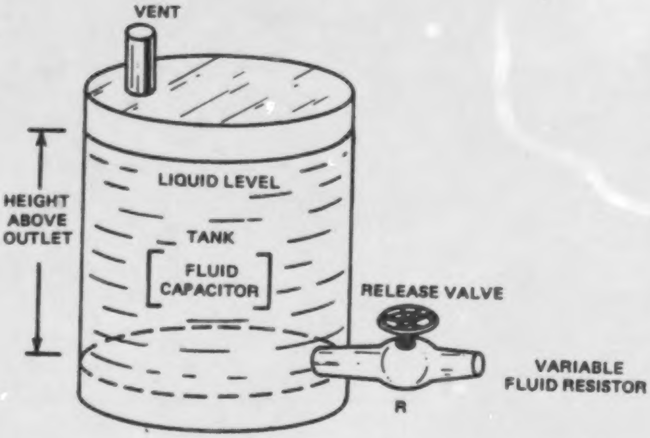


FIG. 1

The fluid capacitance can be defined by Equation 1.

$$C = \frac{V}{\Delta p} \tag{1}$$

where

C = fluid capacitance.
V = volume of fluid stored above the outlet.
Δp = pressure difference across the volume of fluid.

Equation 1 is analogous to the electrical definition of capacitance, which is the electrical charge stored in the capacitor divided by the electrical potential difference across the capacitor. For cylindrical tanks, the volume is the product of the height of the fluid level above the outlet and the cross-sectional area of the tank, as shown in Figure 2 and Equation 2.

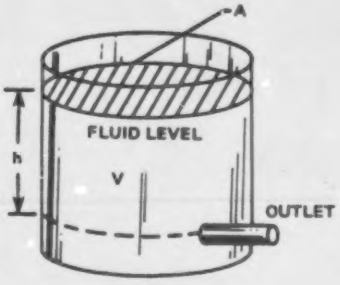


FIG. 2

$$V = Ah \tag{2}$$

where
A = cross-sectional area of tank.
h = height of fluid level above outlet.

Pressure difference between top and bottom level of fluid is given by Equation 3.

$$\Delta p = \rho_w h \tag{3}$$

where
ρ_w = weight density of the fluid = ρg.

The capacitance of a cylindrical tank then is given by Equation 4, derived from Equations 1, 2, and 3.

$$C = \frac{V}{\Delta p} = \frac{(Ah)}{(\rho_w h)}$$

$$C = \frac{A}{\rho_w} \tag{4}$$

EXAMPLE A: FLUID CAPACITANCE OF A CYLINDRICAL STORAGE TANK

Given: In a chemical plant, there is a cylindrical storage tank having a cross-sectional area of 100 m^2 and containing alcohol, which has a mass density of 785 Kgm/m^3.

Find: The capacitance of the storage tank in volume per unit pressure units.

Solution:

$$C = \frac{A}{\rho_w}$$

where
ρ = mass density
g = acceleration of gravity

$$= \frac{A}{\rho g}$$

$$= \frac{100\ m^2}{(785\ Kgm/m^3)(9.8\ m/s^2)}$$

$$C = 0.013\ \frac{m^2}{\frac{Kgm}{m^3} \cdot \frac{m}{s^2}}$$

Dimensionally,

$$\frac{m^2}{\frac{Kgm}{m^3}\cdot\frac{m}{s^2}} = \frac{m^2 \cdot m}{\left[\frac{Kgm\cdot\frac{m}{s^2}\cdot m}{m^3}\right]} = \frac{m^3}{\left[\frac{Newton}{m^2}\right]} = \frac{m^3}{pa}$$

Remember: 1 Pascal (pa) = $\frac{Newton}{m^2}$

Therefore,

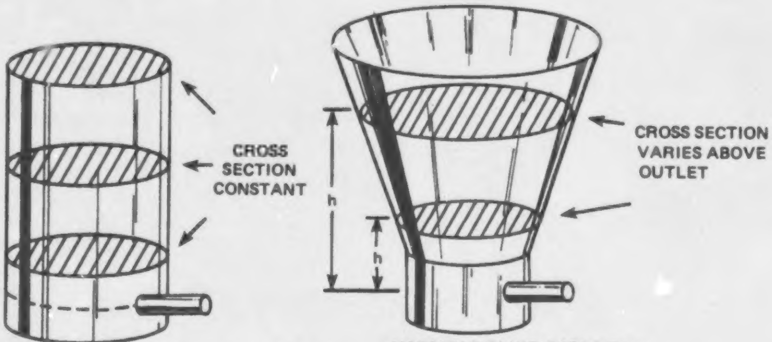
$$C = 0.013\ \frac{m^3}{pa}$$

volume units

pressure units

Equation 4 shows that the capacitance of a cylindrical tank is proportional to the cross-sectional area (surface area of the liquid). Thus, a cylindrical tank with large surface area has greater fluid capacitance than a small-surface-area cylindrical tank containing fluid of the same weight density. Since the cross-sectional area of a tank (surface area of fluid) remains constant as liquid level drops in a cylindrical tank being emptied, such a tank represents a hydraulic capacitance of fixed value. Just as an electrical capacitance does not depend upon the quantity of charge stored at any time, the hydraulic capacitance of a cylindrical tank does not depend upon the quantity of fluid stored at any time.

Tanks whose cross-sectional area varies with the height of fluid level above the outlet, however, have a fluid capacitance that varies as the tank is emptied. Figure 3 illustrates this difference in different types of tanks. This module is concerned only with fixed fluid capacitors; analysis of variable fluid capacitors is beyond the scope of this module.



FIXED FLUID CAPACITOR

VARIABLE FLUID CAPACITOR

FIG. 3

The process of emptying a cylindrical tank, as illustrated in Figure 4, is analogous to the discharge of an electrical capacitor and resistor in series. Concept Module 12-0 presents the equations of discharge for such a resistor-capacitor combination. Similar equations govern the manner in which the volume flow rate of fluid varies with time as a cylindrical tank is emptied. Here the displacementlike quantity (charge) in the electrical system is

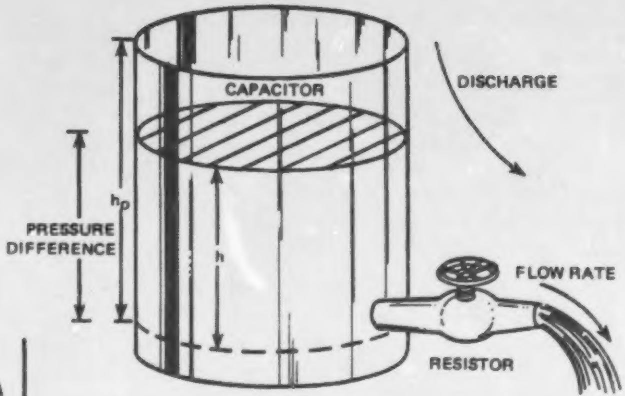


FIG. 4

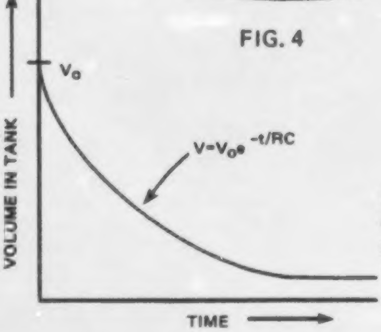


FIG. 5

replaced by the displacementlike quantity (volume) in the fluidal system. The volume remaining in the tank at any time is given by Equation 5 and plotted in Figure 5.

$$V = V_o e^{-t/RC} \qquad (5)$$

where

V = volume in tank at any time.
V_o = maximum volume at $t = 0$.
t = elapsed time since tank began to empty.
R = fluid resistance of the outlet.
C = fluid capacitance of the tank.

Since the volume of a cylindrical tank is directly proportional to the height of the fluid above the outlet, a similar equation exists for the height of the fluid column. Equation 6 gives the height of the liquid at any known time, if the initial height, the resistance, and the capacitance of the tank are known.

$$h = h_o e^{-t/RC} \qquad (6)$$

where

h = height of liquid at any known time.
h_o = initial height of liquid at $t = 0$.
t = any known time.

The product RC is known as the "time constant." The time constant τ_c is the time required for the liquid level to drop to 0.368 h_o. Practically speaking, a tank would be emptied after 5 time constants of elapsed time. Equation 6 is used to solve a problem in Example B.

EXAMPLE 8: FLUID LEVEL IN AN EMPTYING TANK

Given: A water-storage tank has a fluid level 5.00 m above its outlet and a time constant of 2 hours.

Find:

(a) Height of the liquid level after 20 minutes.
(b) Height of the liquid level after 2 hours (one time constant).
(c) Height of the liquid level after 10 hours (five time constants).

Solution: The time constant τ_c = RC = 2 hours (120 min).

(a)
$$
\begin{aligned}
h &= h_o e^{-t/RC} \\
&= (5.00\text{m})e^{-\left(\frac{20\text{ min}}{120\text{ min}}\right)} \\
&= (5.00\text{m})e^{-1/6} \\
&= (5.00\text{m})(0.846) \\
h &= 4.23\text{m}
\end{aligned}
$$

(b)
$$
\begin{aligned}
h &= h_o e^{-t/RC} \\
&= (5.00\text{m})e^{-\left(\frac{2\text{ hr}}{2\text{ hr}}\right)} \\
&= (5.00\text{m})e^{-1} \\
&= (5.00\text{m})(0.368) \\
h &= 1.84\text{m}
\end{aligned}
$$

(c)
$$
\begin{aligned}
h &= h_o e^{-t/RC} \\
&= (5.00\text{m})e^{-\left(\frac{10\ \cancel{\text{hr}}}{2\ \cancel{\text{hr}}}\right)} \\
&= (5.00\text{m})e^{-5} \\
&= (5.00\text{m})(0.0067) \\
h &= 0.0335\text{m}
\end{aligned}
$$

(∿3.3cm -- about empty)

The time constant also is the ratio of the height in the tank at any time to the height change rate, as given by Equation 7.

$$\tau_c = \frac{h \text{ at time } t}{\left(\frac{\Delta h}{\Delta t}\right) \text{at time } t} \tag{7}$$

where

$\frac{\Delta h}{\Delta t}$ = height change rate.

On the graph of Equation 6, shown in Figure 6, the time constant, therefore, is the ratio of the h at any time to the slope of the h-t curve at that time. These quantities are illustrated at two different points on the curve. This equation is useful for determining time constants graphically from an h-t plot, constructed from actual measurements during the emptying of a cylindrical tank.

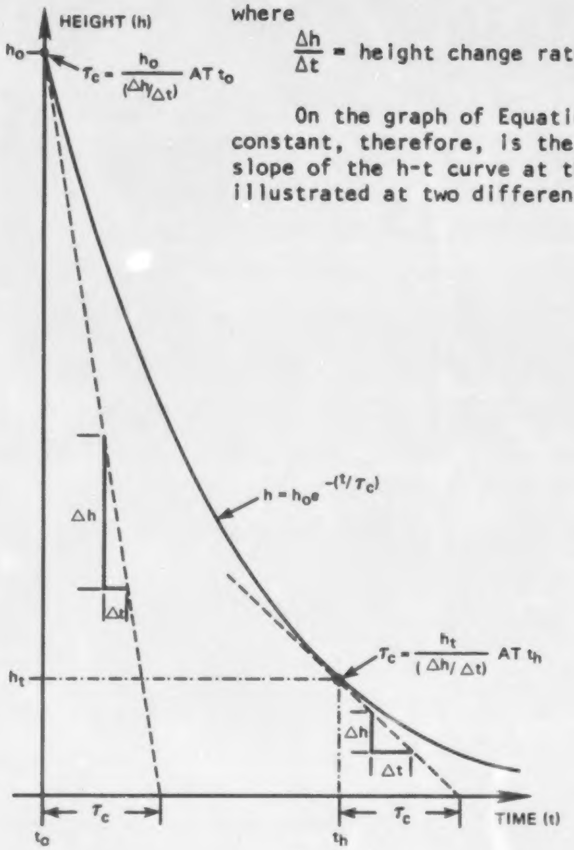


FIG. 6

In order to determine analytically the time constant of a hydraulic, capacitor-resistance, cylindrical-tank system, Equation 6 is solved for τ_c(=RC); and the result is given in Equation 8.

$$\tau_c = \frac{t}{\ln\left(\frac{h_o}{h}\right)} \tag{8}$$

where

h_o = maximum fluid height at t = 0.

h = fluid level at time = t.

t = elapsed time since emptying process began.

EXAMPLE C: FINDING TANK PARAMETERS

Given: A storage tank with water of 10 m initial depth and 50 m^2 cross-sectional area is drained through an outlet at the bottom. After 5 minutes the fluid level has dropped to 9.5 m.

Find:

(a) The time constant of the tank.
(b) The capacitance of the tank.
(c) The outlet resistance.
(d) The time to drain the tank.
(e) The volume flow rate at a fluid level of 5 m above the outlet.

Solution: (a) Use Equation 8:

$$\tau_c = \frac{t}{\ln\left(\frac{h_o}{h}\right)}$$

$$= \frac{(5\ \text{min})}{\ln\left(\frac{10\ \cancel{m}}{9.5\ \cancel{m}}\right)}$$

$$\tau_c = 97.5\ \text{min}$$

(b) The capacitance of the system is given by Equation 4:

$$C = \frac{A}{\rho_w} = \frac{A}{\rho g} \qquad \left[\rho\ \text{water} = 1000\ \text{Kgm/m}^3\right]$$

$$= \frac{(50\ \text{m}^2)}{(1000\ \text{Kg/m}^3)(9.8\ \text{m/s}^2)} = 0.0051\ \frac{\text{m}^3}{\left[\frac{(\text{Kg m/s}^2)}{\text{m}^2}\right]}$$

$$= \frac{\text{N}}{\text{m}^2} = 1\ \text{pa}$$

$$C = 0.0051\ \frac{\text{m}^3}{\text{pa}}$$

(c) Since $\tau_c = RC$ —

$$R = \frac{\tau_c}{C}$$

$$= \frac{\left(97.5\ \cancel{\text{min}} \times \frac{60\ \text{s}}{\cancel{\text{min}}}\right)}{\left(0.0051\ \frac{\text{m}^3}{\text{pa}}\right)}$$

$$R = 1.15 \times 10^6\ \frac{\text{pa}}{\left(\frac{\text{m}^3}{\text{s}}\right)}$$

(d) Practically, $\tau_{empty} = 5\ \tau_c$:

$$\tau_{empty} = 5\ (97.5\ \text{min})$$

$$\tau_{empty} = 487.5\ \text{min}\ (8\ \text{hours}\ 7.5\ \text{min})$$

(e) From Equation 7 —

$$\left(\frac{\Delta h}{\Delta t}\right) = \frac{1}{\tau_c}\ h$$

Multiplication by the cross-sectional area (A) —

$$A\left(\frac{\Delta h}{\Delta t}\right) = \frac{1}{\tau_c}\ Ah$$

$$\frac{\Delta v}{\Delta t} = \frac{1}{\tau_c}\ Ah$$

$$= \frac{1}{(97.5\ \text{min})}\ (50\ \text{m}^2)(5\ \text{m})$$

$$\frac{\Delta v}{\Delta t} = 2.56\ \frac{\text{m}^3}{\text{min}} \qquad \text{(volume flow rate at 5 m)}$$

Notice that since the capacitance of the cylindrical tank is a fixed value, the shape and curvature of the exponential curve of Figure 6 is determined by the value of the resistance to flow. If the resistance to flow of the outlet is small, the exponential graphs will exhibit more curvature, however, if the resistance to flow at the outlet is large, the exponential graph has such a large time constant that the exponential curve has little curvature and appears to be almost linear, as shown in Figure 7.

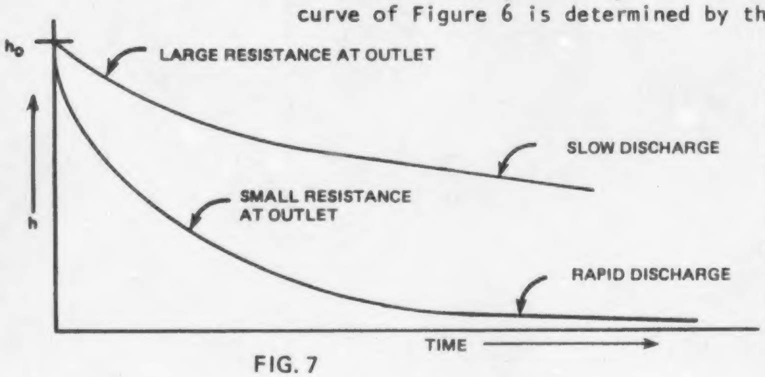


FIG. 7

SUMMARY

The process of emptying cylindrical tanks can be thought of as the discharge of a volume of fluid through a fluid capacitor (the tank) and a resistor (the outlet) in series combination. The product of the fluid resistance and fluid capacitance of the tank is the "time constant" of the hydraulic system. In one time constant, a system discharges to approximately 37 percent of its original volume. After five time constants, the cylindrical storage tanks are practically empty.

RATE OF EMPTYING TANKS

MATERIALS

A large container (preferably transparent).
[Suggestion: Automatic Pipet Rinser found in Chemistry Laboratories or 1 liter Graduated cylinder with hole drilled about 2 cm above the base. Diameter of hole may range from 1/8 to 1/4 inch.]
Stop watch.
Meter stick.
A laboratory sink and water supply.
Rubber stopper to plug orifice.
Felt-tip pen or marking pencil.
Graph paper, 10 squares per inch.

PROCEDURE

1. Arrange equipment as shown in Figure 8.
2. With a felt-tip pen, mark reference level "-0-" at level of center of orifice on side of container (tank).
3. Measure vertically with meter stick from reference "0" level. Mark 5 cm steps vertically as far as possible, as shown in Figure 8.
4. Mark vertical distance of h_{max} (the maximum height available with tank full); record in Data Table 1.
5. Measure inside diameter of tank and diameter of orifice, and record these values in Data Table 1.

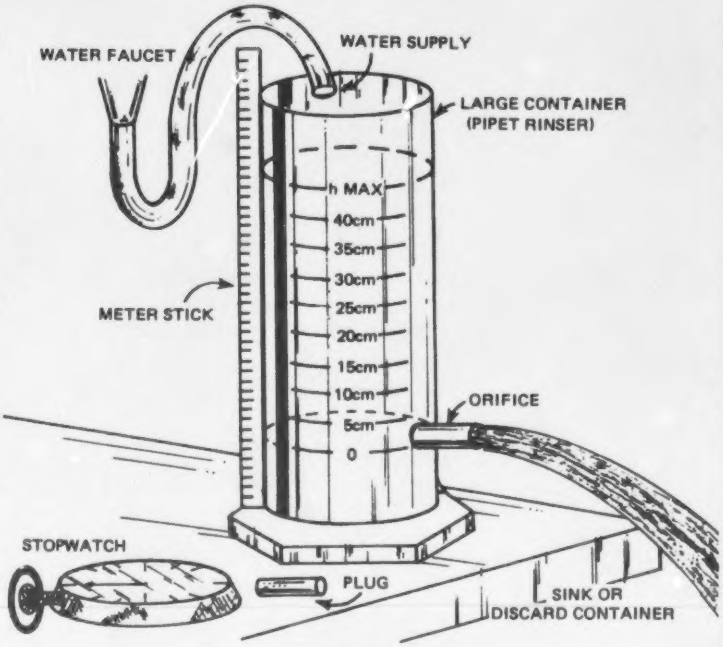

WATER FAUCET
WATER SUPPLY
LARGE CONTAINER (PIPET RINSER)
h MAX
40cm
35cm
30cm
25cm
20cm
15cm
10cm
5cm
0
METER STICK
ORIFICE
STOPWATCH
PLUG
SINK OR DISCARD CONTAINER

FIG. 8

6. Plug orifice tightly, and fill tank until water level is at maximum height. Shut off water supply.

7. Start stop watch when the orifice plug is removed, and measure time required to empty tank. Stop the stop watch the moment orifice ceases to discharge in a continuous flow. Record in Data Table 1 the time required to empty.

8. Repeat procedure of steps 6 and 7, setting liquid level at each successive 5-cm mark marked on side of container and recording in Data Table 1 the corresponding times to empty.

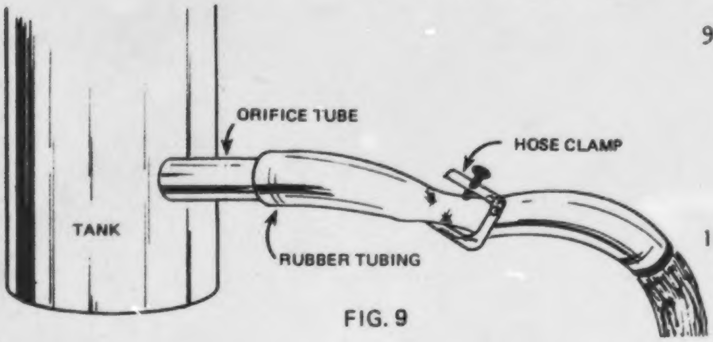


FIG. 9

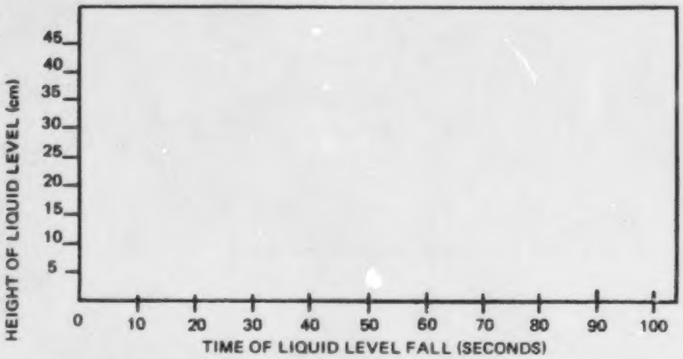


FIG. 10

9. Repeat experiment steps 6, 7, and 8 after increasing resistance to flow by attaching a short piece of rubber tubing with a hose clamp, as shown in Figure 9. Record results in Data Table 2.

10. Calculate time of fall for each level of both trials by subtracting each successive time to empty. Record in the respective columns of Data Tables 1 and 2.

11. On graph paper, plot height-of-liquid-level versus time-of-liquid-level-fall. (Lay out graph as illustrated in Figure 10. Use 10 graph-paper squares per inch.)

12. From graph plotted in Step 11, calculate an average time constant τ_c for the tank with orifice for the whole set of data. Follow graphical procedure outlined in Figure 6 in text.

13. By interpolation from same graph, using $h = 0.368\ h_{max}$, find time constant.

14. Calculate fluid capacitance of container from Equation 4.

15. Using τ_c obtained in step 12 or 13 and fluid capacitance C in step 14, calculate fluid resistance for each orifice used in the experiments.

16. Explain how the laboratory results confirm the principles given in the Discussion section.

DATA TABLE 1. INITIAL RESISTANCE		
HEIGHT OF WATER LEVEL ABOVE ORIFICE (cm)	TIME TO EMPTY (s)	TIME OF FALL OF LIQUID LEVEL (s)
________ (h_{max})	________ (t_{max})	0
40		
35		
30		
25		
20		
15		
10		
5		
2		

Diameter of orifice = ________ cm

Diameter of cylindrical tank = ________ cm

DATA TABLE 2. INCREASED RESISTANCE		
HEIGHT OF WATER LEVEL ABOVE ORIFICE (cm)	TIME TO EMPTY (s)	TIME OF FALL OF LIQUID LEVEL (s)
________ (h_{max})	________ (t_{max})	0
40		
35		
30		
25		
20		
15		
10		
5		
2		

SELF CHECK

1. Define "fluid capacitance" for a cylindrical storage tank.

2. A certain cylindrical water-storage tank has a cross-sectional area of 300 m^2 and an orifice resistance of 3×10^7 pa/(m^3/s). Calculate the time constant of the tank and outlet system.

3. In a tank 10-m tall, the initial fluid level is 8 m above the outlet. The time constant for the system is 20 minutes. What is the level of the fluid above the outlet twenty-six minutes after the outlet is opened?

4. The time constant for a long, cylindrical, chemistry-laboratory buret is 3 seconds. What is the time required to empty the fluid contents out of the bottom outlet?

5. A storage tank of water of 15 m initial depth above the outlet and 75 m^2 cross-sectional area is drained through an outlet at the bottom. After 3 minutes, the level of the surface has dropped to 11 m. Calculate the following:

   a. The time constant of the tank-outlet system.
   b. The capacitance of the tank in m^3/pa.
   c. The outlet resistance in pa/(m^3/s).
   d. The time required for the tank to drain.
   e. The volume flow rate at a fluid level 4 m above the outlet.

6. Find the fluid capacitance of a tank that has a cross-sectional area of 100 m^2 and 5 m of water above the outlet.

12E3

RLC CIRCUITS

In many electronic circuits, it is necessary to deliver electrical power in the form of pulses. For low-energy pulses, resistance-capacitance (RC) circuits may be used. But RC circuits have characteristics that make them undesirable for higher-energy pulses. In all high-energy pulsed circuits and in many low-energy ones, an inductor is added to shape the pulse.

In this module, we shall examine characteristics and applications of resistance-inductance-capacitance (RLC) circuits. The student will determine proper values for critically damped RLC circuits and will examine pulse shapes of overdamped, underdamped and critically damped RLC circuits in the laboratory.

OBJECTIVES

To complete this module, the student should be able to:

- Draw pulse shape from (a) an overdamped RLC circuit (and explain its disadvantages), (b) an underdamped RLC circuit (and explain its disadvantages), (c) a critically damped RLC circuit.
- Given capacitance and resistance of an RLC circuit, calculate inductance necessary for critical damping and pulse duration at one third of maximum height.
- Given necessary equipment, construct RLC circuits and observe pulses from overdamped, underdamped and critically damped circuits.

DISCUSSION

Figure 1 shows the schematic of a typical RLC circuit. In this circuit, the capacitor stores energy that will be delivered to the load resistor R in the form of a pulse. The inductor L is a current limiting device that is used to control shape of the pulse.

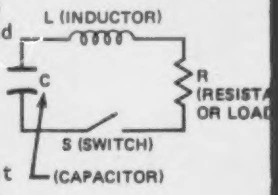


FIG. 1

If inductance is very small, the circuit is said to be overdamped. In this case, circuit performance approximates that of an RC circuit discussed in Module 12E1. Figure 2 shows current through the resistor as a function of time after the switch is closed.

When the switch is closed, current rises rapidly to a peak value of:

$$I = \frac{V}{R}$$

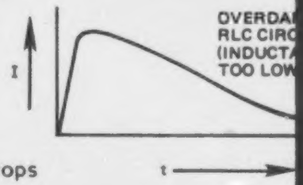


FIG. 2

where V is initial voltage of the capacitor. Current then drops off exponentially as in an RC circuit. The major difficulty with such pulses is the sharp initial rise in current. This rapid change in current produces both thermal and mechanical shocks in load and will greatly decrease useful life of the load resistor. In flashlamp circuits, this sharp current peak is often responsible for lamp burn out and reduces number of shots available from a lamp even if it does not prematurely burn out.

Figure 3 shows the pulse produced by an underdamped RLC circuit. In this case, inductance is larger than the optimum value. Since the inductor resists changes in current, it smooths out the leading edge of the current pulse. When capacitor voltage reaches zero, the inductor tries to maintain current flow at a high level. This causes a reverse charge on the capacitor. At some point, current begins to drop. When it reaches zero, the capacitor has a large reverse charge. This charge causes a current through the load resistance in the opposite direction from the original current. Thus, the circuit oscillates or "rings" until all original energy stored in the capacitor is dissipated. The ringing of such a circuit results in inefficiency of power delivery to the load resistor and may damage energy storage capacitors by rapidly reversing their polarity.

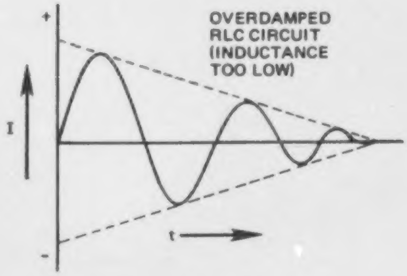


FIG. 3

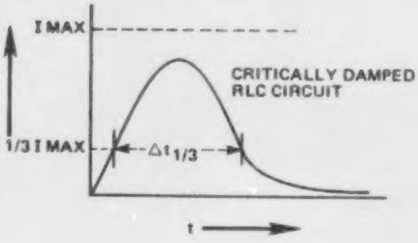


FIG. 4

Figure 4 shows the pulse from a critically damped RLC circuit. In this circuit, value of inductance has been chosen to be large enough to prevent the sharp initial current rise but small enough to prevent "ringing" of the circuit. This delivers energy to the load in the most efficient manner without subjecting either load or energy storage capacitor to unnecessary damaging conditions.

In critically damped RLC circuits, pulse duration is usually measured as full width of the pulse at 1/e of the maximum current. Since 1/e = 0.368, an approximation of 0.333 or 1/3 of maximum height is often used (Figure 4). Pulse duration of such a circuit is given by Equation 1.

$$\text{Pulse duration at } \frac{1}{3} \text{ maximum} = \text{resistance of load} \times \text{capacitance.}$$

$$\text{In symbols:} \quad \Delta t_{1/3} = RC. \qquad (1)$$

For critical damping of an RLC circuit, inductance is given by Equation 2.

$$\text{Inductance} = \frac{\text{square of resistance} \times \text{capacitance}}{4}.$$

$$\text{In symbols:} \quad L = \frac{R^2C}{4}. \qquad (2)$$

Examples A and B demonstrate use of Equations 1 and 2 in solving problems.

Example B represents the most common type of problem encountered. Load resistance and desired pulse duration are usually known. From these values, necessary capacitance and inductance may be determined.

Energy delivered to the load is energy stored in the capacitor as given in Equation 3.

EXAMPLE A: TYPICAL RLC CIRCUIT PROBLEM

Given: An RLC circuit designed to use a 0.4-Ω and a 1000-μf energy storage capacitor.

Find: (a) Pulse duration and (b) inductance necessary for critical damping.

Solution:
(From Equation 1)

$$
\begin{aligned}
\text{(a) } \Delta t_{1/3} &= RC \\
&= (0.4\ \Omega)(1000 \times 10^{-6}\ f) \\
&= 0.4 \times 10^{-3}\ s \\
\therefore \Delta t_{1/3} &= 0.4\ \text{ms.}
\end{aligned}
$$

(From Equation 2)

$$
\begin{aligned}
\text{(b) } L &= \frac{R^2C}{4} \\
&= (0.25)(0.4\ \Omega)^2(1000 \times 10^{-6}\ f) \\
&= 4 \times 10^{-5}\ H \\
L &= 4.0\ \mu H = 40\ \text{microhenries.}
\end{aligned}
$$

EXAMPLE B: TYPICAL RLC CIRCUIT PROBLEM

Given: RLC circuit designed to deliver a 1-ms pulse to a 1-Ω load.

Find: (a) Capacitance necessary and (b) inductance necessary for critical damping.

Solution:

$$
\begin{aligned}
\text{(a) } \Delta t_{1/3} &= RC \\
C &= \frac{\Delta t_{1/3}}{R} \\
&= \frac{10^{-3}\ s}{1\ \Omega} \\
&= 10^{-3}\ f \\
\therefore C &= 1000\ \mu f = 1000\ \text{microfarads.}
\end{aligned}
$$

$$
\begin{aligned}
\text{(b) } L &= \frac{R^2C}{4} \\
&= (0.25)(1\ \Omega)^2(1000 \times 10^{-6}\ f) \\
&= 2.5 \times 10^{-4}\ H \\
\therefore L &= 250\ \mu H = 250\ \text{microhenries.}
\end{aligned}
$$

$$\text{Energy} = \frac{1}{2} \times \text{capacitance} \times \text{square of voltage.}$$

In symbols: $E = \frac{1}{2} CV^2$. (3)

Example C shows use of this equation in solving a problem.

EXAMPLE C: CALCULATION OF ENERGY STORED IN A CAPACITOR

Given: Energy delivered to the load in Example A must be 500 J.

Find: Voltage to which capacitor must be charged.

Solution:
(From Equation 3)

$$E = \frac{CV^2}{2}$$

From which,

$$V^2 = \frac{2E}{C}$$

$$V = \sqrt{\frac{2E}{C}}$$

$E = 500$ J

$C = 1000$ µf.

Substituting,

$$V = \sqrt{\frac{2 \times 500}{1000 \times 10^{-6}}}$$

$$V = \sqrt{10^6 = 10^3}$$

$$\therefore V = 1000 \text{ volts.}$$

Capacitor must be charged to 1000 volts!

RLC circuits are employed when high-energy pulses must be delivered to a load. Common applications involve generation of high-power radar pulses or pulsed magnetic fields for research. A common application involves production of optical pulses from flashlamps. This includes photographic strobes and optically pumped pulsed laser power supplies.

A simplified flashlamp power supply is shown in Figure 5. Resistance of the circuit is lamp resistance after breakdown and is typically between 0.3 Ω and 1.0 Ω. Voltage on the capacitor is sufficient to maintain a current through the lamp but is not high enough to cause electrical breakdown of the gas within the envelope. A trigger pulse of several thousand volts is supplied by a trigger transformer. When this pulse is applied to an external electrode on the lamp, the gas is ionized and lamp resistance drops from a previously very high value to a much lower operating level.

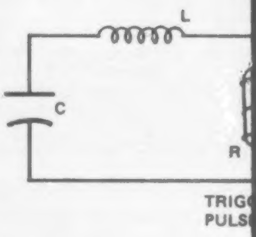


FIG. 5

SUMMARY

RLC circuits are useful for delivering electrical power in pulses quickly and efficiently. Critically damped pulses can be developed by choosing the proper values of R, L and C. High power pulse delivery is often used in radar, flashlamps and pulsed magnetic fields.

LABORATORY

MATERIALS

Oscilloscope
VOM.
Photoflash, flashlamp (National Camera F4407 or equivalent).
Trigger transformer (Archer 272-1146 or equivalent).
550-μF, 400-V capacitor.
0.022-μF, 400-V capacitor.
d.c. power supply, 300-V capability, adjustable.
Push-button switch, mounted.
Clip leads.
50-μH inductor.
200-μH inductor.
50-kΩ, 2-W resistor.
2-MΩ, 1-W resistor.
100-kΩ, 1-W resistor.
0.1-Ω, 2-W resistor or resistance wire to make 0.1-Ω resistor.

PROCEDURE

CAUTION: Dangerous high voltages are present in the circuit used in this experiment. Do not touch components while power is on. Turn off power supply and discharge all capacitors before changing wiring connections.

1. Use clip leads to connect circuit as shown in Figure 6. Turn on power supply; adjust control to obtain a voltage of 300 V as measured with VOM.
2. Place a square of black paper over lamp to reduce amount of room light falling on lamp. Do not look directly at lamp when it is fired.
3. Press push button. If lamp does not fire, have lab instructor check circuit. Allow 15 seconds between shots for lamp to cool.
4. Turn on oscilloscope; set it for 10 V/cm and 0.2 ms/cm. Triggering should be set to give a single sweep when lamp is fired. If you do not understand how to do this, ask instructor for assistance.

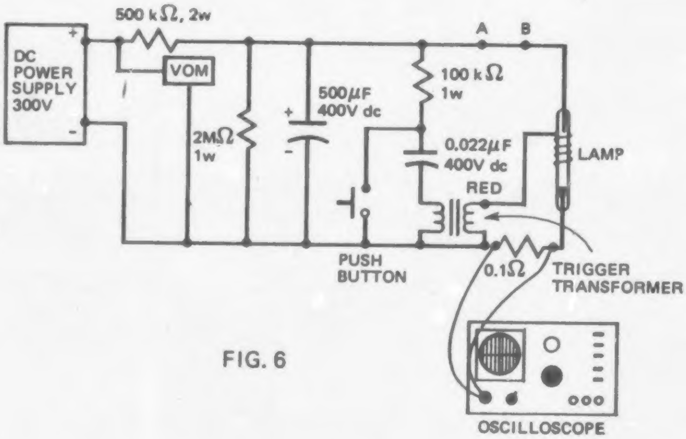

500 kΩ, 2w
A
B
DC POWER SUPPLY 300V
+
-
VOM
2MΩ 1w
500μF 400V dc
+
-
100 kΩ 1w
0.022μF 400V dc
RED
LAMP
PUSH BUTTON
0.1Ω
TRIGGER TRANSFORMER
OSCILLOSCOPE

FIG. 6

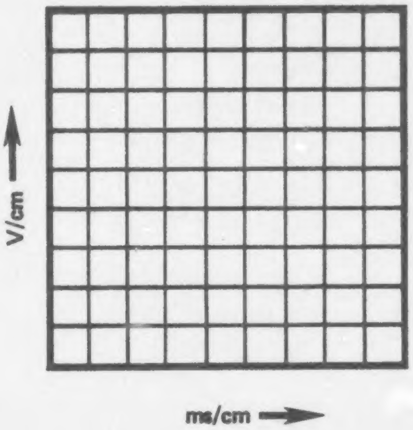
5. Fire lamp several times and observe pulse. Make a sketch of pulse and complete part A of data table.
6. Connect 200-μH inductor between points A and B of Figure 6; repeat above steps. Complete part B of data table. Be sure to note "ringing" frequency.
7. Replace 200-μH inductor with 50-μH inductor; repeat again. Complete part C of data table.

NOTE: The 0.1-Ω resistor in the circuit is called a "current-viewing" resistor and is present for measurement only. It is not normally included in a flashlamp circuit.

DATA TABLE

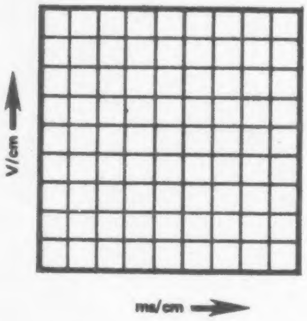
Part A: Overdamped RLC circuit

Sketch pulse produced with zero inductance between A and B. Include pulse duration.



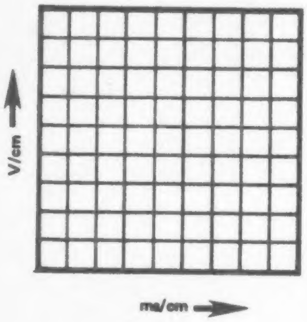
Part B: Underdamped RLC circuit

Sketch pulse produced with an inductance of 200 μH between A and B. Include ringing frequency.



Part C: Critically damped RLC circuit

Sketch pulse produced with an inductance of 50 μH between A and B. Include pulse duration.



SELF CHECK

1. Draw pulse shape from an overdamped RLC circuit; explain its disadvantages. Why is circuit overdamped?
2. Draw pulse shape from an underdamped RLC circuit; explain its disadvantages. Why is circuit underdamped?
3. Draw pulse shape from a critically damped RLC circuit.
4. 600 J of energy is to be delivered to a 0.5-Ω load with a pulse duration 0.5 ms. Find:

   Capacitance necessary.
   Inductance necessary for critical damping.
   Voltage to which energy storage capacitor must be charged.

5. Use pulse duration measured for the critically damped circuit in experiment and capacitance of the circuit; calculate resistance of the load.
6. An RLC circuit is overdamped. What is the maximum value of the current flowing in the circuit if the initial voltage is 1 volt and the resistor has a value of 100Ω.
7. What resistance should be placed into an LC circuit in order that the resulting RLC circuit be critically damped if L = 10 mH and C = 20 μf?
8. How much energy does a 10 μf capacitor store when the capacitor is charged to a voltage of 2.5 volts?
9. What capacitance should be added to an RL circuit in order that the resulting RLC circuit have a time constant of 0.1 seconds if R = 1000Ω.

RESPONSE TIME OF THERMOCOUPLES

A thermocouple consists of a pair of wires of dissimilar metals joined together at one end, with opposite ends connected to a voltage-measuring device. When temperature of the junction between the wires changes relative to the other ends, a voltage is generated. Voltage varies in relation to temperature change and can be used to monitor temperature.

Time is required for thermal energy (heat) to move through materials by thermal conduction. If one end of a metal spoon is placed in hot water, after some length of time, the other end will become warm. The rate of such heating effects can be discussed in terms of convenient time intervals, referred to as "time constants." In particular, a temperature-measuring device monitoring temperature of a certain sample takes time to respond to a change in the sample's temperature. Its response time is governed by the time required for thermal conduction of heat from the sample to the measuring device.

This module discusses heat in transient situations; that is, in situations in which temperature is changing rapidly. Under these conditions, response time of a temperature-measuring device becomes important. Response time of a thermocouple depends upon its size, its material, and its bonding to the material under study. Small, thin-film thermocouples, for example, have short response times (microseconds).

In the laboratory, the student will measure response times for a conventional thermocouple and a thin-film thermocouple and will use a thermocouple to measure thermal diffusivity of a sample of metal.

OBJECTIVES

Upon completion of this module, the student should be able to:

- Define thermal diffusivity and time constant of a thermocouple and give an equation for each.
- Calculate the response time for a given thermocouple.
- Calculate and plot the temperature indicated by a thermocouple that is subjected to a sudden change of temperature, given necessary information about the thermocouple's properties.
- Compare and contrast the importance of mass, area, and bonding to material as these factors affect the relative response time of conventional and thin-film thermocouples.
- Given the appropriate equipment, measure the response time of a thin-film thermocouple and of a conventional thermocouple; and measure thermal diffusivity with a thermocouple.

DISCUSSION

Heat flow is dependent on thermal conductivity (k) and specific heat (c) of the substance. In fact, the heat flow or heating rate of a substance is inversely proportional to the specific heat per unit volume, which is equal to ρc, where ρ is the material density. An important factor for determining heat flow is $k/\rho c$. This factor has the dimensions of cm^2/s, characteristic of a diffusion coefficient, and therefore has been given the descriptive term "thermal diffusivity" (to identify it as the diffusion coefficient of heat). We shall refer to it as k_d.

For changing or transient conditions of heat flow, the term $k/\rho c$ is important. The factor is significant because it determines how rapidly a material will accept and conduct thermal energy. Values of thermal diffusivity k_d for a few metals are given in Table 1.

TABLE 1. THERMAL DIFFUSIVITY OF COMMON METALS

Metal	Thermal Diffusivity (k_d)
Aluminum	0.86 cm^2/s
Brass	0.34
Copper	1.14
Nickel	0.155
Steel (carbon)	0.121
Steel (stainless)	0.040

When a certain amount of heat energy is applied to a material, the heat energy moves (propagates) through the material in a finite time. In many practical cases, propagation can be defined in terms of a time constant. The changing temperature T of the substance being heated as a function of time t follows a relation analogous to Equation 1.

$$T = T_f - (T_f - T_i)\, e^{-t/t_c} \qquad (1)$$

where

T_f = final temperature.

T_i = initial temperature.

t_c = time constant of the particular substance.

When $t = 0$, $e^{-t/t_c} = 1$, and $T = T_i$. When t becomes very large, $e^{-t/t_c} \cong 0$, and $T = T_f$.

Equation 1 supposes that heat energy is delivered beginning at time $t = 0$ to a material at temperature T_i. When temperature is being measured at a distance from the position at which heat energy is being supplied to a material with thermal diffusivity k_d, the thermal time constant often is given approximately by Equation 2.

$$t_c = d^2/k_d \tag{2}$$

where

d = distance from point of temperature measurement to point at which heat energy is applied.

k_d = thermal diffusivity.

This approximate relation is not applicable in all cases; still, it defines a measure of the thermal time constant.

Obviously, all the above concepts hold equally well whether the substance is heated or cooled; thus, heat flow into or away from a given substance (and use of a time constant) have equal validity.

Example A illustrates the use of Equations 1 and 2 in studying temperature versus time for a heated aluminum bar.

EXAMPLE A: GRAPHICAL SOLUTION OF BAR-END TEMPERATURE

Given: An aluminum bar 10 cm long is initially at temperature 20°C. At time $t = 0$, heat is supplied to one end; heating continues until the bar reaches a uniform steady state temperature of 80°C.

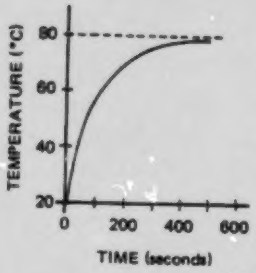
Find: Use Equations 1 and 2 to make a graph of temperature versus time for the far end of the bar.

Solution: From Equation 2 and Table 1:

$t_c = 10^2/0.86 = 116.3$ seconds

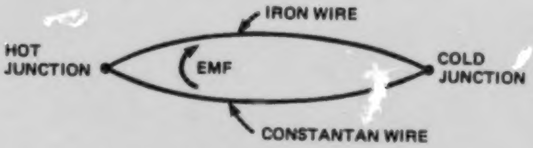
Substituting in Equation 1 gives:

$T = 80 - 60\ e^{-t/116.3}$

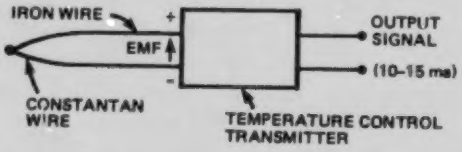


The above drawing gives results obtained by graphing this equation. An exponential approach to the final temperature of 80°C is observable. A thermocouple monitoring temperature of the end of the bar would show an output of a similar form.

Propagation of heat across an interface of two different materials is also important. A thermocouple (Figure 1) can be attached to a solid sample, or it can be immersed in a fluid. Heat must be transferred between the material (either solid or fluid) and the thermocouple. An exact description of heat transfer across such a boundary is very complicated. One need only note that heat transfer across a boundary involves a factor called "surface heat transfer coefficient" (k_h) This coefficient defines how much heat energy is transferred across a unit area of the boundary during a unit of time for a unit of temperature


IRON WIRE
HOT JUNCTION
EMF
COLD JUNCTION
CONSTANTAN WIRE

(b) A TYPICAL THERMOCOUPLE AND TRANSMITTER


IRON WIRE
+
EMF
-
OUTPUT SIGNAL
(10-15 ma)
CONSTANTAN WIRE
TEMPERATURE CONTROL TRANSMITTER

(a) AN IRON-CONSTANTAN THERMOCOUPLE

FIG. 1

difference. The surface coefficient (k_h) has units of calories/cm^2-s-C°. For two solids, the surface coefficient will depend on how well the surfaces are in contact. For a solid and a liquid, the value depends on thickness of a boundary layer of liquid, across which temperatures vary. For instance, two solids separated by a thin layer of grease 0.1 mm thick will have a surface heat transfer coefficient of approximately 0.01 calories/cm^2-s-C°.

Surface heat transfer coefficients are important in determining how heat is transferred from a solid to a thermocouple.

The actual, in-use response time of a thermocouple will be affected by two factors:

(a) Time taken for heat to be conducted from its point of application in the material to the position of the thermocouple in the material. (Note that this time lapse has nothing to do with the properties of the thermocouple.)

(b) Time taken for heat to be transferred from a material to the thermocouple and for the thermocouple to respond. (Note that this is what usually is referred to as the thermocouple response time.)

In many cases (a) above may be estimated by using concepts described in relation to Equations 1 and 2, and Example A. Often this time will be very short. For example, if a thermocouple is plunged suddenly into a liquid of a different temperature, the time represented by factor (a) will be negligible, and response of a thermocouple to an instantaneous change of temperature is governed essentially by factor (b).

For an instantaneous change in temperature from T_1 to T_2, temperature T indicated by the thermocouple will be given by Equation 3.

$$(T - T_2) = (T_1 - T_2)\, e^{-t/t_c} \qquad (3)$$

Thus, response will be governed by the time constant t_c.

EXAMPLE B: CALCULATION OF TEMPERATURE WITH VARYING TIME

Given: A thermocouple at temperature 20°C is plunged into a bath of boiling water. The thermocouple has a response time of 0.1 second.

Find: Temperature the thermocouple indicates after 0.1, 0.2 and 1.0 seconds.

Solution: Substitute into Equation 3 to obtain the temperature T as a function of time t.

$t_c = 0.1$ s

$T_1 = 20$°C (initial temperature)

$T_2 = 100$°C (boiling point of water)

$T = 100 - 80\, e^{-t/0.1}$

When $t = 0.1$ s, $T = 100 - 80\, e^{-1} = 70.6$C°

When $t = 0.2$ s, $T = 100 - 80\, e^{-2} = 89.2$C°

When $t = 1$ s, $T = 100 - 80\, e^{-10} = 99.996$C°

The thermal time constant t_c of a thermocouple is given by Equation 4.

$$t_c = \frac{mc}{k_h A} \tag{4}$$

where

m = mass of thermocouple.
c = specific heat of thermocouple.
k_h = coefficient of surface heat transfer from material to thermocouple.
A = area of contact between thermocouple and material.

Example C illustrates the use of Equation 4 in calculating the time constant for a thermocouple of given properties and given surface heat transfer coefficient (k_h).

EXAMPLE C: CALCULATION OF THERMOCOUPLE TIME CONSTANT

Given: A thermocouple of mass 0.002 g and specific heat 0.08 cal/g - C° is attached to a sample through a layer of grease 0.1 mm thick. The area of contact between thermocouple and sample is 0.02 cm^2.

Find: Time constant of the thermocouple.

Solution: Use Equation 4, with the value for k_h of 0.01 cal/cm^2-s - C°. Substituting in Equation 4 produces this result:

$$t_c = \frac{mc}{Ak_h}$$

$$\frac{(0.002)(0.08)}{(0.01)(0.02)}$$

$$t_c = 0.8 \text{ second}$$

Usually, time constants of the thermocouple should be small, in order that the thermocouple can monitor changing temperature without a large time delay. Equation 4 indicates four important factors for reducing the time constant:

- Mass of a thermocouple should be as small as possible.
- The thermocouple should be made of material with low specific heat, if possible. Often this will not be possible, because choice of materials will be governed by thermoelectric response. Materials will be chosen to give a large voltage change for a given temperature change, rather than for low values of specific heat.
- Area of contact between thermocouple and sample should be as large as possible.
- Thermocouple should be in good thermal contact with sample; this will lead to a large value of surface heat transfer coefficient (k_h).

In some cases, a compromise between response time and ruggedness of the thermocouple must be made. If a thermocouple is to be used in an environment in which vibration and mechanical shock are present, the necessity may arise to increase its mass and modify the attachment to make installation more durable.

These changes can decrease the response time; thus, a low-mass device for rapid response must be balanced against a larger device for improved tolerance to mechanical shock and vibration.

Efforts to attain a fast thermocouple response center on the use of thin metallic films, evaporated on the sample. Reference to Equation 4 reveals why thin-film thermocouples can offer faster responses.

A conventional thermocouple often is made of two wires, bonded together, with a small drop of melted and refrozen metal. Thermocouples are attached to a sample with some good heat-conducting material, such as certain greases or adhesives, to make them stick. The drop of metal sets a lower limit to the mass. It also determines the area of contact between thermocouple and sample. A bonding material also limits thermal conduction from sample to thermocouple. According to Equation 4, all these factors influence the time constant. Again, use of thin metal film improves response time.

In a thin-film thermocouple, metals are evaporated to form layers between 10^{-5} and 10^{-4} cm thick, reducing the mass to a minimum. They are deposited directly on the material; thus, there is no need for an intermediate material to attach the thermocouple. Thermal contact with the sample is maximized, therefore, and surface heat transfer coefficient is high. A thin film also may be spread over a larger surface area than the dimension of a conventional thermocouple.

In thin-film thermocouples, time constants as low as one microsecond have been achieved. Thus, a thin-film thermocouple can offer truly rapid response in situations of changing temperature. In contrast, conventional thermocouples (two thin wires with a small bead of metal at the junction) may have response speeds of a few tenths of a second — five orders of magnitude larger!

Investigations of methods of producing conventional thermocouple junctions have been carried out (reference Green and Hunt). Thin wires around 0.01-inch diameter were used. Wires were chromel (nickel with 10% chromium) and alumel (nickel with 3% manganese, 2% aluminum and 1% silicon). Wires were spot welded with a resistance welder under carefully-controlled conditions to form a bead no larger than the original wire diameter. Even with these precautions, response speeds were found to be about 10 milliseconds.

Workers using evaporated thin films of copper and constantan (copper with between 35% and 50% nickel) were able to obtain response times of less than one microsecond. The thermocouples were formed by the evaporation of metals in a vacuum and by allowing the metal vapor to condense on glass substrates. Thickness of the layers was about 1.2×10^{-5} cm. Response time was measured by heating the junctions with a pulsed laser (reference Chopra, Bahl and Randlett).

More recently, fast thermocouples have been formed by a thin tube of metal forced over a fine wire, with 0.0005-in thick insulation between elements. Contact between tube and wire is formed by the evaporation of a thin film of metal on the end of the structure. This variation of commercially-available thin-film thermocouples also has response times as small as one microsecond.

Use of thin-film thermocouples involves extra expense and labor, compared to conventional thermocouples. In many applications, increased speed of response is not needed, and conventional thermocouples are satisfactory. For example, in monitoring a furnace, the temperature will not change significantly on a microsecond time scale. Conventional thermocouples with time-constant ranges from tenths of a second, or perhaps seconds, will be fast enough to give an accurate measure of changing temperature.

Thermocouples are routinely used in industry to monitor furnace or oven temperatures under conditions of heating or cooling. Thermocouples are an important element in a feedback system that stabilizes temperature. In this application, the thermocouple must respond rapidly to varying temperature.

Thermocouples are employed to measure temperature of flowing fluids under conditions of variable temperature. Once again, if temperature of the fluid changes rapidly, the thermocouple must sense the change rapidly.

Thermocouples are used in nuclear reactors, where speed can be critical. As conditions in a reactor change, the reactor core may heat up. To keep temperature within acceptable limits, sensors must measure the temperature change rapidly, in order that responsive control systems can reduce heating.

Fast-response thermocouples are used in a variety of specialized scientific research and engineering applications, such as the measurement of temperatures inside gun barrels and the recording of heating effects produced by high-power pulsed lasers. They are also used in measuring heatup of rocket nose cones.

SUMMARY

The response time of thermocouples is unimportant in many temperature-measurement applications; however, in other applications — such as nuclear reactors, temperature measurement of swift fluid flow, and aerospace equipment — a very fast response is required. The thin-film thermocouple provides the faster response time of the two major types (the thin-film and the wire types). Response time depends upon the following characteristics of a thermocouple: (1) size, (2) material, and (3) bonding to the material under study.

RESPONSE TIME OF CONVENTIONAL AND THIN-FILM THERMOCOUPLES

MATERIALS

Dewar flask of liquid nitrogen (or a beaker of water and ice if liquid nitrogen is not available).
Wires for a thermocouple (e.g., copper and constantan, or other metals if available).
Soldering iron and solder.
Oscilloscope (should have vertical sensitivity of a few millivolts/division).
Thin-film thermocouple. (Medtherm Corp., P. O. Box 412, Huntsville, Alabama 35804, or equivalent).
Electrical leads and wires as needed.
Oscilloscope camera and Polaroid film, if available.

PROCEDURE

In this experiment, the student will measure the response time of thermocouples by plunging them in a liquid at a different temperature and observing a voltage pulse. Results for a conventional thermocouple are to be compared with a thin-film thermocouple.

1. Fill a small dewar flask with liquid nitrogen. (If not available, use a beaker of water and ice.) Liquid nitrogen, if available, will give easier-to-measure, larger changes of voltage.

2. Prepare a thermocouple (copper and constantan or other appropriate metals). Twist ends of wires together; solder them. Make solder bead as small as possible. Attach opposite ends to oscilloscope input.

3. Plunge thermocouple into cold liquid. The sudden change of temperature will generate a change in voltage, which can be viewed on oscilloscope. Several trials may be required to adjust sweep speed and vertical sensitivity. After each trial, allow thermocouple to warm to room temperature. Oscilloscope trace should appear something like drawing in Figure 2. Total change in voltage will

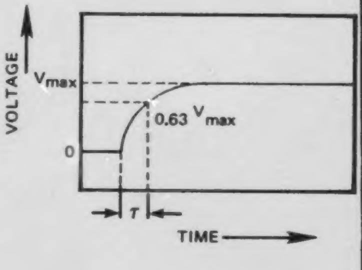


FIG. 2

be a few millivolts if liquid nitrogen is used; less if water and ice are used.

4. On oscilloscope trace, measure time at which voltage change has reached 63% of its final value. Simplify this step with a scope camera and Polaroid film, if available. (Remember from discussion in Module 12-0 that a parameter has reached 63% of its final value when time elapsed is equal to time constant. This measurement thus gives the constant directly. The 63% level is indicated at right of Figure 2.)

5. Repeat steps 3 and 4 with a thin-film thermocouple. Compare time constants obtained with two types of thermocouples.

6. Why does the thin-film thermocouple have a faster response?

THERMAL DIFFUSIVITY OF A METAL

MATERIALS

Dewar flask of liquid nitrogen (or a beaker of water and ice if liquid nitrogen is not available).
Wires for a thermocouple (e.g., copper and constantan, or other metals if available); same or similar type as used in first experiment.
Strip chart recorder. (Suppliers include Honeywell, Inc., Test Instruments Division, Denver, Colorado. Honeywell, model Electronik 19 has a 0.1-millivolt scale and would be suitable. Esterline-Angus, Indianapolis, Indiana, is another possible supplier.).
Stirring rod.
Adhesive (e.g., General Radio cement).
Stainless steel rod about 10 cm long (rods of other metals are acceptable).
Clamp and stand to support rod.
Ruler.

PROCEDURE

In this experiment, the student will use a thermocouple under conditions of changing temperature to determine thermal diffusivity of a metal.

1. Attach thermocouple to one end of metal rod, using adhesive.

2. Attach thermocouple leads to strip chart recorder.

3. Turn on recorder; place other end of rod into liquid nitrogen (or beaker of water and ice). Apparatus is sketched in Figure 3. Support rod with a suitable clamp.

4. Observe change of voltage on strip chart. Stir liquid with stirrer in order that cold liquid is kept in contact with ends of rod. Metal will cool gradually. As far end of rod cools, voltage indicated by chart recorder will change. Under conditions described here, a few hundred seconds may be required for voltage to reach its final steady value. Use a slow chart speed.

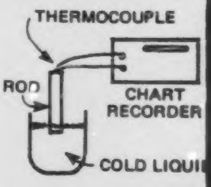


FIG. 3

5. From strip chart, determine time t_c at which voltage has changed 63% from initial value to final value.

6. Analyze results, using the following equation for approximate voltage V as a function of time t:

$$V = V_f + (V_i - V_f)\, e^{-k_d \pi^2 t/4\ell^2}$$

where

V_i = initial value of voltage.
V_f = final value of voltage.
k_d = thermal diffusivity of metal.
ℓ = length of rod (measured with ruler).

This equation represents the approximate temperature of one end of a rod of material, when the other end is suddenly placed in contact with a well-stirred fluid at time zero. Note that at time zero, the exponential factor is one, and —

$$V = V_f + (V_i - V_f) = V_i.$$

At large values of time t, the exponential factor approaches zero, and —

$$V = V_f.$$

From this equation, the value of the time constant t_c as determined from step 5 is —

$$t_c = 4\ell^2/k_d\pi^2$$

Thermal diffusivity k_d is derived as follows:

$$k_d = 4\ell^2/\pi^2 t_c$$

Use measured values to calculate k_d. Compare the results to values in Table I, and explain any difference in values.

SELF CHECK

1. Define thermal diffusivity, using an equation. Also using an equation, define time constant of a thermocouple.

2. A thermocouple has a mass of 0.006 g, a specific heat of 0.15 cal/g degree, a contact area of 0.01 cm^2 and a surface heat transfer coefficient of 0.03 cal/cm^2-s-°C. What is time constant of this thermocouple?

3. The thermocouple of problem 2 is at 20°C. It is suddenly plunged into liquid nitrogen (temperature -196°C). Graphically plot temperature indicated by the thermocouple as a function of time. Calculate indicated temperature at several values of time, indicating initial rapid change and approach to a steady final temperature.

4. Compare and contrast mass, area, and bonding to material as these factors affect relative response time of conventional and thin-film thermocouples.

5. Estimate the thermal diffusivity of a material that has a time constant of 10 minutes if the distance from the point of temperature measurement to the point at which the heat energy is applied is 10 cm.

6. What is the heat capacity (mc) of a thermocouple that has a time constant of 1 second if the area of contact between the thermocouple and the material is 0.01 cm^2 and k_h = 0.01 cal/cm^2·s·C°?

7. A thermocouple at temperature 30°C is placed into a pan of water that is at a temperature of 95°C. What is the time constant of the thermocouple if after being in the pan of hot water for 0.1 s the temperature indicated by the thermocouple is 92°C?

12T2

TIME CONSTANTS OF HEATING AND COOLING

Control of the rate of heating and cooling is an important function in many industrial situations. For example, the cooling rate of hot glass for telescope mirror blanks must be controlled during the manufacturing process. If the blanks were cooled at too fast a rate, the outside layers of glass would be cool, while the layers underneath would remain hot. The resulting thermal stress would cause cracks to occur, damaging thousands of dollars worth of glass work. Time constants associated with the response of temperature-measuring devices, such as liquid-filled thermometers, are important factors to consider in the design of an accurate temperature-measuring system. This module describes time constants for heating and cooling and relates the time constants to the rate of temperature changes of a body as it heats or cools. In the laboratory, the student will determine how different surface characteristics of a hot body affect the time constant and the rate at which it cools.

OBJECTIVES

Upon completion of this module, the student should be able to:

- Explain how time constants are used to measure the rate at which temperature approaches an equilibrium state.
- Given an initial temperature, ambient temperature, and time constant for a material body, determine its temperature at any time.
- Given the appropriate equipment, calculate time constants of materials with different thermal resistances.

DISCUSSION

When heat flows into a body, and no mechanical work is done by the body, all the added energy shows up as stored internal energy, as evidenced by a rise in temperature of the body. A release of this stored energy to its surroundings is evidenced by a decrease in the body temperature. See Figure 1. The direction of heat flow is governed by the law that states, "Heat travels naturally from a hot body (at high temperature) to a cooler body (at low temperature)." When a body is at the same temperature as its surroundings, no heat is transferred; and the body is said to be in "equilibrium" with its surroundings. The larger the temperature difference between two different bodies, the faster the heat flow between them.

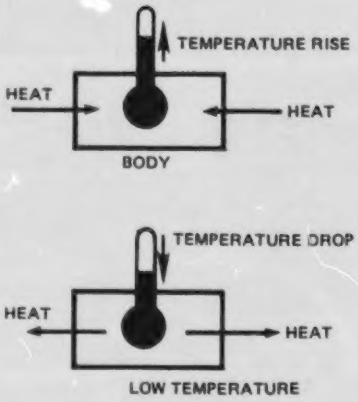


FIG. 1

Heat is transferred by three methods — conduction, convection, and radiation. Conduction is a method of heat transfer that occurs on a molecular scale, and is the predominant method of heat transfer through solids. Convection is a method of heat transfer that occurs by movement of hot and cold masses of material. Convection is an important method of heat transfer in fluids. Radiation is a method of heat transfer in which heat is transferred by means of electromagnetic radiation. The rate at which heat is transferred by this method is strongly dependent upon the surface characteristics of the body. Black-colored objects radiate heat at the fastest rate. Automotive racers use car engines of increasing displacement, causing problems of excessive heat. Blackened engine blocks, rocker arm covers, and oil pans aid in getting the heat out of the automobile engine at the fastest rate. Car radiators are painted black for the same reason. Silver, chrome, or mirror-surfaced bodies radiate heat at a slower rate. A thermos bottle uses this characteristic to reduce the rate at which its contents lose or gain heat. Most often these methods of heat transfer occur together, so that total heat transfer occurs as a result of a combination of conduction, convection, and radiation. An example of such a combined effect is shown in Figure 2.

Often, in practical applications, the total rate of heat flow is difficult to describe; however, the total rate of heat flow is dependent upon the difference in temperature between the source of heat and a surrounding fluid, which maintains a constant ambient temperature (Figure 2). The total heat flow rate also is affected by the total combined thermal resistance of conduction, convection, and radiation effects.

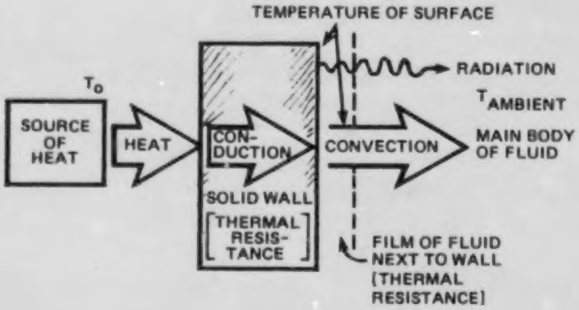


FIG. 2

THE RATE OF HEATING

The process of heating a cool body to a higher temperature occurs in our daily lives frequently. The doctor inserts a thermometer (cool body) in the mouth of a patient (hot surroundings) and waits three or four minutes while the thermometer comes to the temperature of the patient's body. A housewife places a turkey (cool body) in the oven (hot surroundings) and then waits for the meat to reach cooking temperature. In each case, the hot surroundings maintain a constant temperature while the cooler body gradually increases its temperature.

Consider a thermometer initially at T_o (50°F) being plunged into a bath at temperature T_B (100°F). The temperature is recorded at intervals of 5 seconds, as shown in Figure 3. The temperature of the thermometer at any time is given by Equation 1.

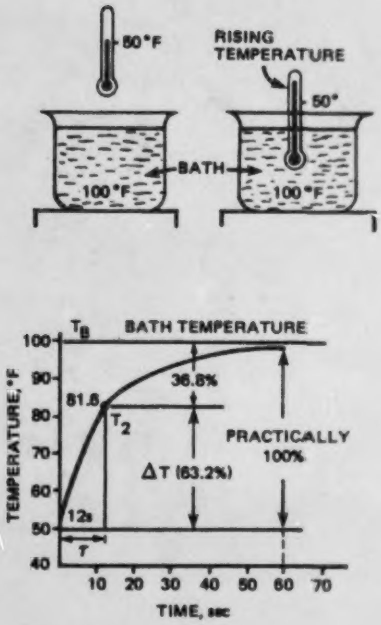


FIG. 3

$$T = T_o + (T_B - T_o)[1 - e^{-t/\tau}] \qquad (1)$$

where

T = temperature of thermometer at time t.

T_o = temperature of thermometer at time t = 0.

T_B = constant temperature of hot surroundings (bath temperature).

e = exponential base = 2.71828.

τ = time constant of thermometer.

As discussed in Concept Module 12-0, the time constant (naperian base) is the time required for a body to reach 63.2% of the difference $(T_B - T_o)$. For the example shown in Figure 3, the time constant of the thermometer is 12 seconds. After the 12 seconds, the temperature has risen to 81.6°F. After 5 time constants (5 x 12 seconds = 60 seconds), the thermometer, for all practical purposes, has risen to 100°F. The time constants associated with half-life processes are not normally used in thermal work.

The time constant also can be defined as the ratio of the temperature difference $(T_B - T)$ and the rate of temperature change at that temperature. It is given by the relation indicated in Equation 2.

$$\tau = \frac{(T_B - T)}{\left(\frac{\Delta T}{\Delta t}\right)} \qquad (2)$$

where

T_B = temperature of hot surroundings.
T = temperature at time t.
$(\Delta T/\Delta t)$ = rate of temperature change at T (slope of tangent line at point t, T on graph).

Equation 2 can be rearranged to solve for rate of temperature change, as indicated by Equation 3.

$$\left(\frac{\Delta T}{\Delta t}\right) = \frac{T_B - T}{\tau} \qquad (3)$$

Equations 1, 2, and 3 can be applied to any thermal body placed in hotter surroundings of constant temperature. These equations are used to solve a problem in Example A.

EXAMPLE A: HEATING OF A BODY TO A CONSTANT TEMPERATURE OF ITS SURROUNDINGS

Given: A thermometer initially at 50°F is plunged into a hot-water bath of 200°F. The system time constant is 12 seconds.

Find:
(a) The temperature after one time constant.
(b) The time required for the thermometer to be practically the same temperature as the bath.
(c) The initial rate of temperature change at 50°F.

Solution:

(a)
$$T = T_o + (T_B - T_o)[1 - e^{-t/\tau}]$$
$$= (50°\text{F}) + (200°\text{F} - 50°\text{F})[1 - e^{-\left(\frac{12\ \text{s}}{12\ \text{s}}\right)}]$$
$$= 50°\text{F} + (150°\text{F})[1 - e^{-1}]$$
$$T = 144.8°\text{F}$$

(b) Time for thermal equilibrium = 5 x time constant
= (5)(12 s)
= 60 s

(c) Solve for $(\Delta T/\Delta t)$, using Equation 3:
$$\left(\frac{\Delta T}{\Delta t}\right) = \frac{T_B - T}{\tau}$$
$$= \frac{(200°\text{F} - 50°\text{F})}{12\ \text{s}}$$
$$\left(\frac{\Delta T}{\Delta t}\right) = 12.5\ \frac{\text{F}°}{\text{s}}$$

Regardless of the temperature of the bath, the thermometer will achieve 63% of the difference between the initial temperature of the thermometer and the temperature of bath in one time constant. After 5 time constants, the thermometer almost will achieve equilibrium, regardless of the temperature of the hot surroundings. Of course, the value of the time constant itself depends upon the bath temperature.

THE RATE OF COOLING OF A HOT BODY

The process of cooling a hot body in surroundings of constant cooler temperature occurs in everyday circumstances, such as the heat-treating of metal parts (quenching), the freezing of foods, and the response of temperature-sensing instruments.

Consider a thermometer which is initially in a hot water bath of 170°F and then is withdrawn to still air at 70°F. The decreasing temperature of the thermometer is recorded every 50 seconds, as illustrated in Figure 4. In such a case, the temperature is given by Equation 4.

$$T = T_{AMB} + [T_O - T_{AMB}]e^{-t/\tau} \qquad (4)$$

where

T_{AMB} = ambient temperature.

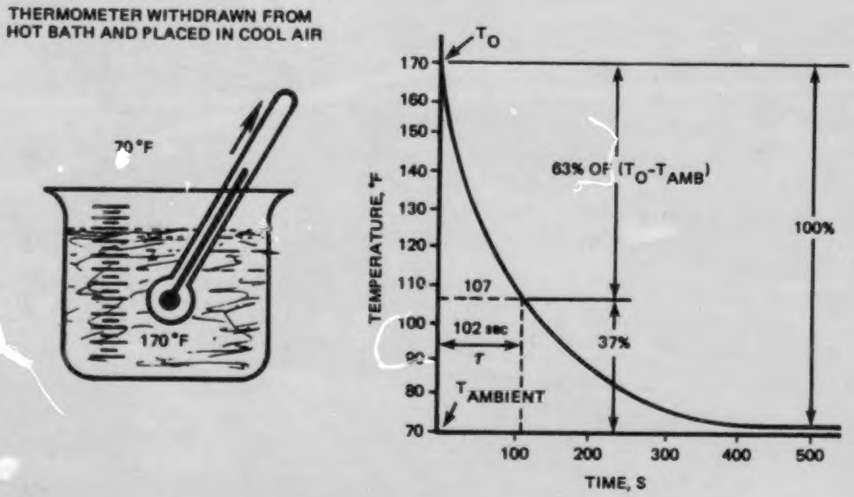


FIG. 4

The time constant will be the time required for the temperature of the body to cool to 63% of (T_o - T_{AMB}). The time constant also is related to the rate of temperature change by Equation 5.

$$\tau = \frac{(T - T_{AMB})}{(\frac{\Delta T}{\Delta t})} \qquad (5)$$

The time required for the hot body to reach ambient temperature is again equal to 5 time constants. In the example in Figure 4, the time constant is 102 seconds. The temperature drops to 107°F after one time constant. After 510 seconds the thermometer essentially reaches the ambient temperature of 70°F.

Equations 4 and 5 are employed to solve a sample problem in Example B.

EXAMPLE B: COOLING OF A BODY TO THE CONSTANT TEMPERATURE OF ITS SURROUNDINGS

Given: A thermometer is immersed in a 120°F hot water bath and is withdrawn and left in still air at 70°F. The thermometer has a time constant of 102 seconds.

Find:
(a) Temperature after one time constant.
(b) Time required for the thermometer to be practically at still-air temperature.
(c) Initial rate of temperature change at 120°F.

Solution:

(a)
$$T = T_{AMB} + [T_o - T_{AMB}]e^{-t/\tau}$$
$$= (70°F) + [120°F - 70°F]e^{-(\frac{\cancel{102\ s}}{\cancel{102\ s}})}$$
$$= 70°F + (50°F)\ e^{-1}$$
$$T = 88.4°F$$

(b)
$$T\ \text{equilibrium} = 5\ \tau$$
$$= 5\ (102\ s)$$
$$T\ \text{equilibrium} = 510\ s$$

(c)
$$\frac{\Delta T}{\Delta t} = \frac{T - T_{AMB}}{\tau}$$
$$= \frac{(120°F - 70°F)}{102\ s}$$
$$\frac{\Delta T}{\Delta t} = 0.49\ \frac{F°}{s}$$

Notice that no matter what the initial temperature of the body is, the body essentially will reach equilibrium in 5 time constants.

FACTORS AFFECTING THE TIME CONSTANTS FOR HEATING AND COOLING BODIES

The value of the time constant for a given body depends upon many factors. The thermal resistance of the material of the body, the specific heat of the material, the surface area of the body exposed to the surroundings, and the mass of the body all affect the rate at which heat passes into, or out of, a body. Another factor that affects the rate of temperature change of a body is the state of motion of the fluid surroundings; for example, a thermometer in still air has a larger time constant than one in moving air.

SUMMARY

The transient processes of heating and cooling a body in a constant-temperature, ambient environment are described by use of an important factor called a "time constant." Specific equations involving temperature, time, and the time constant aid in the modeling of the thermal behavior of a heating or cooling body.

THERMAL TIME CONSTANTS

MATERIALS

3 brake fluid cans of same size and materials (two painted mirror-finish chrome or silver, one painted flat black).
3 rubber stoppers to fit opening in cans (stoppers bored for thermometer).
1 asbestos or cork sheet 1/4" thick as insulating jacket for one of silvered cans.
3 thermometers (0 to 100°C, 1C°).
1 small bottle glycerin (thermometer lubricant).
3 beakers (500 ml volume).
1 beaker tongs.
Wristwatch or clock.
1 room thermometer.

PROCEDURE

1. Arrange brake fluid cans as illustrated in Figure 5. Make sure that they are separated from each other by about 50 cm and that they are supported on insulating paper towels.

2. Heat a large quantity of water to boiling temperature or, at least, above 95°C.

3. While water is heating, check arrangement in Figure 5 to ensure that the thermometers are adjusted about 2 centimeters above bottoms of cans.

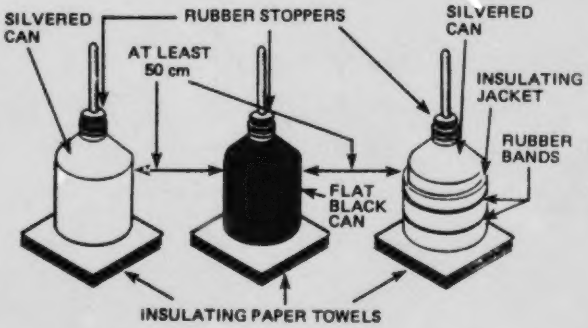


FIG. 5

4. When water has achieved its high temperature, fill each can with an equal volume (300-400 ml) of hot water. Insert the thermometers and rubber stoppers as shown in Figure 5. Notice the rise in temperature of the thermometer.

5. Read the room thermometer, and record this ambient temperature in the Data Table.

6. After thermometers have reached their maximum temperature (in about 3 or 4 minutes), begin recording temperature of each can every two minutes for the first 10 minutes and then once every 10 minutes for 2 1/2 hours. Record time and temperature in Data Table.

7. Plot temperature vs. time curve for each can on the same graph (i.e., three curves on same graph). Identify each curve on graph with proper can.

8. Does the black can cool at a faster rate than the silvered can? If so, why?

9. Which can, the silvered or insulated silvered, cools at a faster rate? Why?

10. From the graph, which can has the largest time constant? The smallest time constant?

11. Determine the time constant for each can from the curve.

12. Based upon the time constants, compute the time required for the cans to reach room temperature.

DATA TABLE

TIME (MIN)	TEMPERATURE OF SILVERED CAN (°C)	TEMPERATURE OF BLACK CAN (°C)	TEMPERATURE OF INSULATED CAN (°C)
0			
2			
4			
6			
8			
10			
20			
30			
40			
50			
60			
70			
80			
90			

100			
110			
120			
130			
140			
150			
ROOM TEMPERATURE (AMBIENT) = ______________ °C			

SELF CHECK

1. Thermometer A has a time constant of 8 seconds. Thermometer B has a time constant of 6 seconds.

   (a) Which thermometer reaches equilibrium quicker?
   (b) For the same initial temperature and hot environment temperature, which thermometer has the larger initial rate of temperature change?

2. Assuming that 4 minutes is about the time required for a doctor to get an accurate temperature measurement of a patient (equilibrium of thermometer and body temperature), what is the time constant of the thermometer used?

3. A chrome-constantan thermocouple has a time constant of 0.7 second, and is initially at a temperature of 80°F. It is plunged into a hot oil bath of 300°F. What is the temperature of the thermocouple after 0.5 second?

4. A 200-lb aluminum engine block cools from 150°F to 100°F in 15 min. Assume that the temperature drops exponentially with time. Surrounding air is 70°F. What is the time constant of the block? What is the temperature after 20 minutes?

5. A small thermocouple cools from 100°C to room temperature in about 35 seconds. If the thermocouple is heated in a flame to 500°C, what is its temperature 14 seconds after the flame is removed? Room temperature is 20°C.

6. A body cools at a rate of 0.25 C°/s when its temperature is 20°C above the ambient temperature. What is the cooling rate when the temperature is 60°C above its surroundings?

RADIATION

APPLICATION MODULES

Page
13L1* Inverse Square Law — Spreading of Light 497
13L2* The Optical Spectrum 509
13L3 Light Sources 522
13L4* Lasers 531
13L5 Stroboscopes — Light Pulses 540
13L6* Optical Filters — Absorption of Light 547
13L7* Specular and Diffuse Reflection of Light 558
13L8* Image Formation with Mirrors 570
13L9* Prisms — Refraction and Dispersion of Light 583
13L10* Lenses — Focusing and Spreading of Light 594
13L11* Interference — Testing Flat Surfaces 606
13L12* Diffraction Grating — Measuring the Wavelength of Light 615
13L13* Polarization — Glare Reduction 625
13L15 Spectroscopy — Light Analysis 638
13N3* Radioactivity 649

* Identifies concept support modules that introduce new physics.

INVERSE SQUARE LAW — SPREADING OF LIGHT

A reader, needing more light, moves closer to the light source. On cold nights, a camper sits closer to the fire. Faint sounds are heard more easily by placing the ear near the source of the sound. These are examples of the inverse-square law of radiation. All radiation moves outward from its source radially in straight lines. Thus, energy of radiation spreads over an ever-increasing area as it travels farther from the source.

This module presents the inverse-square law as it applies to radiant energy (light). Measurements made on an optical system in the laboratory will be used to verify the inverse-square law.

OBJECTIVES

Upon completion of this module, the student should be able to:

- Define:

  Irradiance.
  Solid angle.
  Radiant intensity.

- Write the equation for the inverse-square law of light and define each term in the equation.

- Given any two of the following quantities, calculate the third:

  Radiant intensity of a point source of light.
  Distance from light source to point of observation.
  Irradiance at point of observation.

- Given proper equipment, measure the irradiance produced by a point source of light at several distances from the source. Draw a graph of irradiance versus distance and calculate the radiant intensity of the source.

DISCUSSION

All radiation spreads as it leaves its point source. The discussion of the inverse-square law presented in this module is based upon light radiation, although the same principle and equations apply to radio waves, sound, nuclear radiation and all other types of radiation.

Figure 1 shows light radiating from a point source. The light emitted into the volume bounded by the four lines (A, B, C and D of the figure) continues to travel outward. At distance d this light energy passes through the cross-sectional area A. At 2d, the same energy spreads to cover area 4A. At 3d, it covers 9A. Figure 1 demonstrates that the area through which energy passes increases with the square of the distance from the source. This rate of increase forms the basis of the inverse-square law.

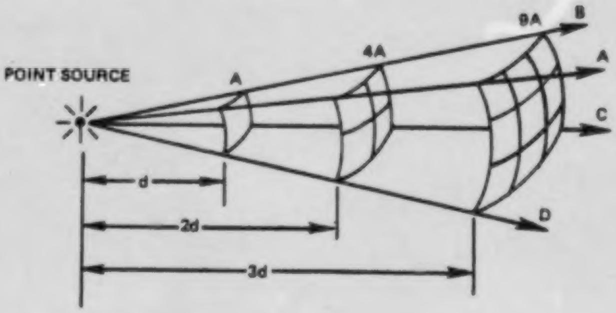


FIG. 1

Before a detailed discussion of the inverse-square law can be presented, several concepts used in the measurement of light must be introduced. The first of these is irradiance, a measure of the power per unit area striking a surface. It is the power divided by area, or power per unit area, as given in Equation 1.

$$E = \frac{P}{A} \qquad (1)$$

where

E = irradiance.
P = power.
A = area.

Example A illustrates the use of this equation in solving a problem.

EXAMPLE A: IRRADIANCE ON A SURFACE

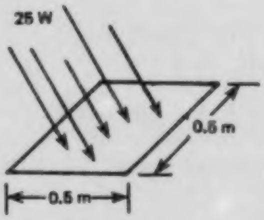
Given: 25W (watts) of light falls on one side of a square of 0.5 meter on each side.

Find: Irradiance at surface.

Solution:

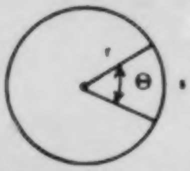
$$A = d^2 = (0.5\ \text{m})^2$$
$$= 0.25\ \text{m}^2$$

$$E = \frac{P}{A} = \frac{25\ \text{W}}{0.25\ \text{m}^2}$$
$$= 100\ \text{watt/meter}^2.$$

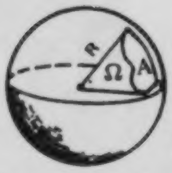


Note that the units of irradiance are "power divided by area," usually expressed in W/m^2 or W/cm^2.

The second concept to be considered is the solid angle illustrated in Figure 2. In plane geometry (Figure 2a), an angle is defined as "that part of the plane lying between two lines that meet at a point." Such an angle may be measured in degress or radians, where 2π radians equals 360 degrees. The angle in radians is defined by Equation 2.



(a) PLANE ANGLE



(b) SOLID ANGLE

FIG. 2

$$\theta = \frac{s}{r} \qquad (2)$$

where

θ = angle in radians.
s = length of arc subtended by angle θ.
r = radius of circle.

Figure 2b shows the analogy in solid geometry. Solid angle, Ω, is defined as "the surface area subtended at center of sphere divided by the square of the radius of sphere," as given in Equation 3.

$$\Omega = \frac{A}{R^2} \qquad (3)$$

where

Ω = solid angle in steradians.
A = area of sphere intercepted by angle Ω.
R = radius of sphere.

6 Solid angles are measured in steradians (sr). Like the radian, the steradian is a dimensionless unit. There are 4π steradians in a complete sphere. Example B illustrates Equation 3.

EXAMPLE B: CALCULATION OF SOLID ANGLE

Given: A circular surface 10 cm in diameter, viewed from a distance of 1 m.

Find: Solid angle subtended by the surface at the eye.

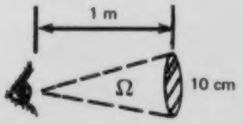
Solution:

$$A = \pi r^2 = (3.14) \times (5\ \text{cm})^2$$

$$= 78.5\ \text{cm}^2$$

$$\Omega = \frac{A}{R^2} = \frac{78.5\ \text{cm}^2}{(100\ \text{cm})^2}$$

$$\Omega = 0.00785\ \text{sr}$$



Note: The circular surface has been assumed to be approximately equal to the spherical surface intercepted by Ω with radius of 1 m.

The concept of a solid angle is useful in describing the way light is radiated from a source. Figure 3 shows the beam from a flashlight. At distance R, the beam illuminates area A with power P. The solid angle of the beam is very nearly given by $\Omega = A/R^2$. The radiant intensity of a light source is defined as the "power per unit solid angle radiated by the source," as given by Equation 4.

$$I = \frac{P}{\Omega} \qquad (4)$$

where

I = radiant intensity.
P = power.
Ω = solid angle.

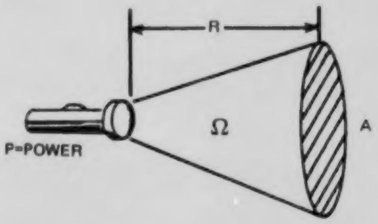


FIG. 3

Example C demonstrates the use of this equation in solving a problem.

EXAMPLE C: RADIANT INTENSITY OF A LIGHT SOURCE

Given: In Figure 3, the flashlight produces 10 W of optical power, $R = 2$ m and $A = 0.5$ m^2.

Find: Radiant intensity of flashlight.

Solution:

$$\Omega = \frac{A}{R^2} = \frac{0.5\ \text{m}^2}{(2\ \text{m})^2}$$

$$= 0.125\ \text{sr}$$

$$I = \frac{P}{\Omega} = \frac{10\ \text{W}}{0.125\ \text{sr}}$$

$$I = 80\ \frac{\text{W}}{\text{sr}}$$

Note: radiant intensity is the function of a light source; it does not depend upon the distance from a source. For instance, if I (in Example C) is recalculated for $R = 4$m, Ω is shown to remain unchanged; therefore, radiant intensity I remains unchanged.

On the other hand, irradiance depends upon source-to-surface distance and decreases as distance from source increases. Irradiance, at any distance, R, from a point source of total power, P, is given by the inverse-square law (Equation 5).

$$E = \frac{P}{4\pi R^2} \qquad (5)$$

where

E = irradiance of a surface a distance R from a point source.
P = total power of point source.
R = distance from point source to surface (point of measurement).

Example D illustrates use of the inverse-square law in solving a problem.

EXAMPLE D: IRRADIANCE E AT DISTANCE R FROM POINT SOURCE

Given: A 500 W electrical arc, treated as a point source, emitting light equally in all directions.

Find: Irradiance at (a) 25 cm, (b) 50 cm and (c) 1 m.

Solution:

(a) $E_a = \frac{P}{4\pi R^2} = \frac{500\ W}{(4) \times (3.14) \times (25\ cm)^2}$

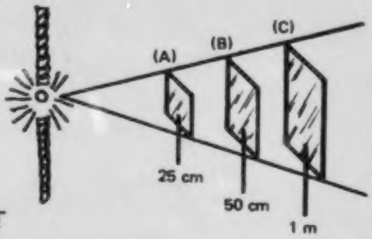
$E_a = 6.37 \times 10^{-2}\ \frac{W}{cm^2}$

(b) $E_b = \frac{500\ W}{(4) \times (3.14) \times (50\ cm)^2}$

$E_b = 1.59 \times 10^{-2}\ \frac{W}{cm^2}$

(c) $E_c = \frac{500\ W}{(4) \times (3.14) \times (100\ cm)^2}$

$E_c = 3.98 \times 10^{-3}\ \frac{W}{cm^2}$



Note: Multiplying distance by 4 (from 25 cm to 100 cm) reduces irradiance by a factor of 16 or 4^2.

Equation 6 presents an alterate form of the inverse-square law.

$$E = \frac{I}{R^2} \qquad (6)$$

where

E = irradiance due to a point source.
I = radiant intensity of the point source.
R = distance from source to point of measurement.

This equation combines Equations 4 and 5.

Examples E and F provide some practice in the use of this equation in solving problems.

EXAMPLE E: IRRADIANCE CALCULATION

Given: Light source with radiant intensity of 250 W/sr.

Find: Irradiance at 10 m.

Solution:

$$E = \frac{I}{R^2}$$

$$= (250\ \mathrm{W/sr}) \div (10\ \mathrm{m})^2$$

$$E = 2.5\ \mathrm{W/m^2}$$

Note: In Example E the unit "steradian" (sr) is dropped from the calculation; as noted previously, a steradian represents a ratio and, therefore, has no dimensions.

EXAMPLE F: RADIANT INTENSITY CALCULATION

Given: At a distance of 2 m from a source, 3.6 mW falls on an area of 0.5 cm^2.

Find: Radiant intensity of source.

Solution: (Use Equation 6).

$$\text{(a)}\quad E = \frac{P}{A} = \frac{3.6 \times 10^{-3}\ \mathrm{W}}{0.5\ \mathrm{cm^2}}$$

$$= 7.2 \times 10^{-3}\ \mathrm{W/cm^2}$$

$$I = ER^2 = (7.2 \times 10^{-3}\ \mathrm{W/cm^2})(200\ \mathrm{cm})^2$$

$$= 288\ \mathrm{W/sr}$$

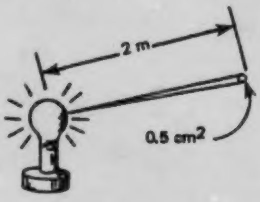
(Use Equation 4).

$$\text{(b)}\quad \Omega = \frac{A}{R^2} = \frac{0.5\ \mathrm{cm^2}}{(200\ \mathrm{cm})^2}$$

$$= 1.25 \times 10^{-5}\ \mathrm{sr}$$

$$I = \frac{P}{\Omega} = \frac{3.6 \times 10^{-3}\ \mathrm{W}}{1.25 \times 10^{-5}\ \mathrm{sr}}$$

$$= 288\ \mathrm{W/sr}$$



SUMMARY

Two quantities (irradiance and radiant intensity) are of major importance in describing the distribution of light energy as it travels from a light source and strikes a surface. Irradiance is a measure of the power per unit area striking a surface. Radiant intensity is a measure of the power per solid angle leaving a source. The inverse-square law describes the rate of decrease in irradiance as the distance from the source increases. If the distance is doubled, the irradiance drops by a factor of four (2^2). If the distance is tripled, the irradiance drops by a factor of nine (3^2), and so on. This relationship applies to all forms of radiation which leave a point source uniformly in all directions.

DEMONSTRATION OF THE INVERSE SQUARE LAW

MATERIALS

Clear incandescent bulb, 10 W or less (flashlight with reflector removed will work).
Optical power meter.
Meter stick.

PROCEDURE

1. If power meter has a limiting aperture, remove it to allow entire detector area to be illuminated.

2. Measure dimesion of detector surface; record in Data Table.

3. Place light bulb and detector as shown in Figure 4 so that bulb filament and detector surface are at the same height and are 10 cm apart.

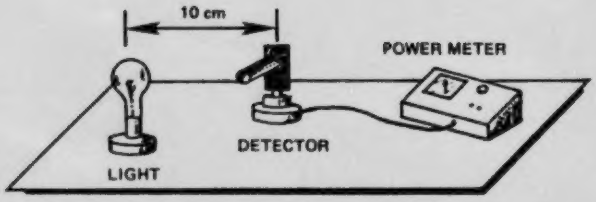


FIG. 4

4. Set power meter to least sensitive scale; turn on power meter and incandescent bulb; measure power and record in Data Table. Change power meter to more sensitive scale, if necessary, but do not drive needle past full-scale reading.

5. Move detector away from source, in 10 cm steps, to 1 m; measure and record power at each distance. Change power meter to more sensitive scale as necessary.

6. Calculate irradiance at each distance (power in milliwatts received divided by detector area in square centimeters).

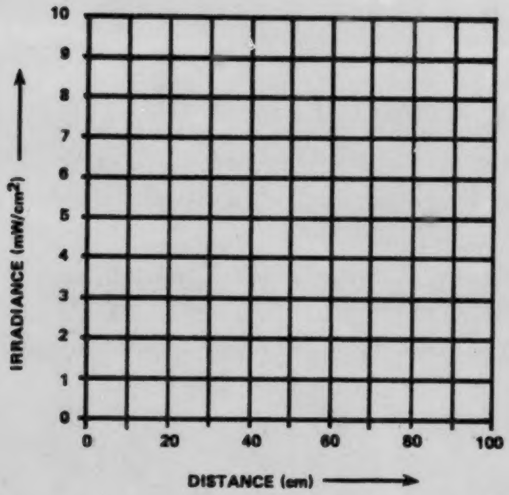
7. Calculate radiant intensity as measured from each distance (use Equation 6).

8. Plot irradiance versus distance on graph paper provided.

9. Explain how graph illustrates inverse-square law of light.

DATA TABLE		
Distance (cm)	Irradiance at Detector $\left(\frac{mW}{cm^2}\right)$	Radiant Intensity of Light Bulb (W/sr)
10		
20		
30		
40		
50		
60		
70		
80		
90		
100		

Detector area = ____ cm²



SELF CHECK

1. Define:

   Irradiance.
   Solid angle.
   Radiant intensity.

2. Write equation for inverse-square law of light; define each term in equation.

3. If radiant intensity of a flashlight is 5 W/sr, calculate its irradiance at 25 m.

4. If radiant intensity of an arc lamp is 150 W/sr and irradiance at a detector is 0.25 mW/cm^2, calculate distance from light source to detector.

5. If irradiance 15 m from a light source is 0.03 mW/cm^2, calculate irradiance at 20 m.

6. Find the area of a sphere intercepted by a solid angle of $\pi/2$ steradians if the radius of the sphere is 3 m.

7. Find the radiant intensity of a source if at a distance of 1.5 m from the source a 5 m^2 area is delivered a power of 20 watts.

13L2

THE OPTICAL SPECTRUM

The term "optics" is often thought of as referring to devices used within the visible portion of the electromagnetic spectrum. Optics and optical instruments are actually employed over a broad spectral range, including not only visible light radiation, but also the adjacent ultraviolet and infrared. The optical spectrum portion is the more important one to us biologically, providing heat, energy for photosynthesis, and the basis for sight itself. This module presents the basic properties and several applications which involve the optical portion of the electromagnetic spectrum. In the laboratory the student will utilize a diffraction grating to examine the visible spectrum.

OBJECTIVES

Upon completion of this module, the student should be able to:

- Label the infrared, visible and ultraviolet portions of the electromagnetic spectrum; supply the wavelength and frequency scales, given a blank chart of the optical spectrum.
- Define the following wavelength units:

    Micron.
    Nanometer.
    Angstrom.

- Given one of the following quantities for a photon, calculate the other two:

    Wavelength.
    Frequency.
    Energy.

- Explain briefly the origin of each of the following types of spectra:

  Line.
  Band.
  Continuous.

- State one application of each of the following regions of the optical spectrum:

  Visible.
  Ultraviolet.
  Infrared.

- Given the appropriate equipment, measure the wavelengths of several lines in a line spectrum.

DISCUSSION

THE OPTICAL SPECTRUM

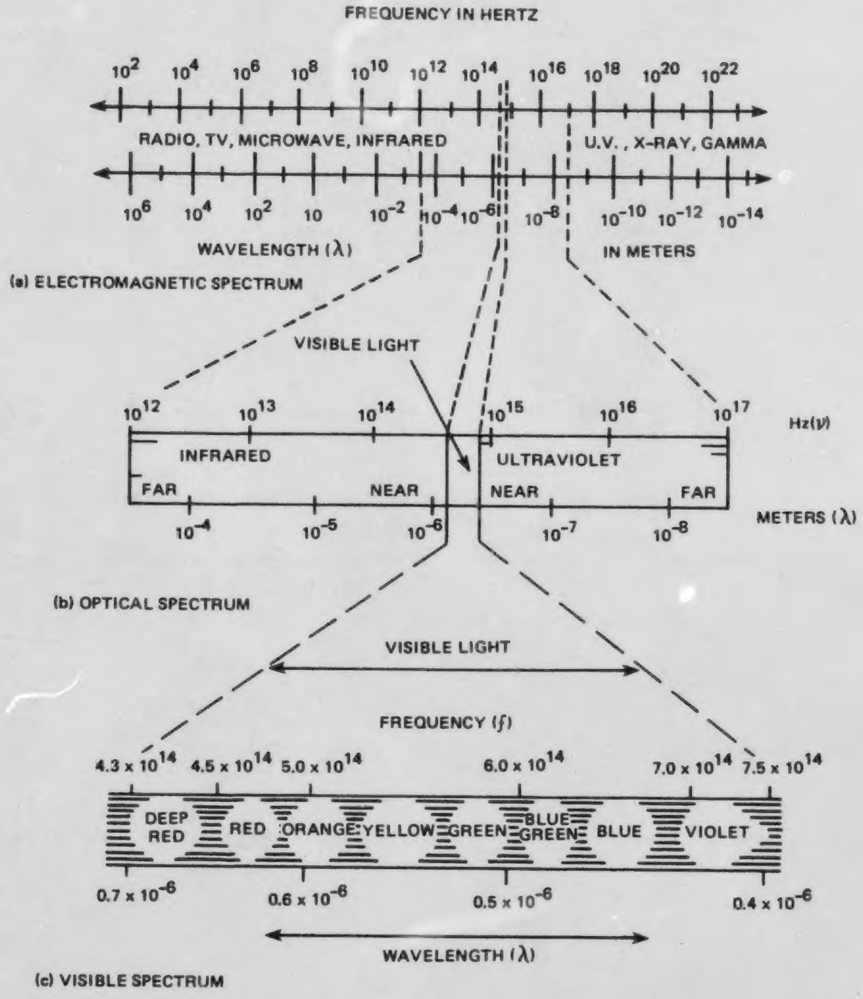

FREQUENCY IN HERTZ
10^2 10^4 10^6 10^8 10^10 10^12 10^14 10^16 10^18 10^20 10^22
RADIO, TV, MICROWAVE, INFRARED
U.V., X-RAY, GAMMA
10^6 10^4 10^2 10 10^-2 10^-4 10^-6 10^-8 10^-10 10^-12 10^-14
WAVELENGTH (λ)
IN METERS
(a) ELECTROMAGNETIC SPECTRUM
VISIBLE LIGHT
10^12 10^13 10^14 10^15 10^16 10^17
Hz(ν)
INFRARED
ULTRAVIOLET
FAR
NEAR
NEAR
FAR
10^-4 10^-5 10^-6 10^-7 10^-8
METERS (λ)
(b) OPTICAL SPECTRUM
VISIBLE LIGHT
FREQUENCY (f)
4.3 x 10^14 4.5 x 10^14 5.0 x 10^14 6.0 x 10^14 7.0 x 10^14 7.5 x 10^14
DEEP RED
RED
ORANGE
YELLOW
GREEN
BLUE GREEN
BLUE
VIOLET
0.7 x 10^-6 0.6 x 10^-6 0.5 x 10^-6 0.4 x 10^-6
WAVELENGTH (λ)
(c) VISIBLE SPECTRUM

FIG. 1

Figure 1a shows the electromagnetic spectrum and the location of the optical spectrum. The optical spectrum (figure 1b) is composed of visible light, the longer-wavelength infrared (IR) light and the shorter-wavelength ultraviolet (UV) light. The wavelength range of most interest in this module is the visible spectrum (Figure 1c) and those wavelengths closest to the visible - called the near-infrared and the near-ultraviolet.

Table 1 lists three length units commonly used to express the wavelengths in the optical spectrum. Angstroms (10^{-10} m) are normally used for the far ultraviolet, for X-rays, and sometimes for the near ultraviolet and visible, although the nanometer (10^{-9} m) is more common for the longer wavelengths. Infrared wavelengths are usually expressed in microns (10^{-6} m). Example A shows the conversion of a wavelength given in meters into each of these units.

TABLE 1. WAVELENGTH UNITS

Name	Symbol	Definition
Angstrom	Å	10^{-10} m
Nanometer	nm	10^{-9} m
Micron	μ	10^{-6} m

EXAMPLE A: CONVERSION OF WAVELENGTH UNITS

Given: The wavelength of a photon is 6.33×10^{-7} m.

Find: Wavelength in (a) angstroms, (b) namometers and (c) microns.

Solution:

(a) $(6.33 \times 10^{-7}\ \text{m}) \times \left(\frac{10^{10}\ \text{Å}}{1\ \text{m}}\right) = 6330\ \text{Å}$

(b) $(6.33 \times 10^{-7}\ \text{m}) \times \left(\frac{10^{9}\ \text{nm}}{1\ \text{m}}\right) = 633\ \text{nm}$

(c) $(6.33 \times 10^{-7}\ \text{m}) \times \left(\frac{10^{6}\ \mu}{1\ \text{m}}\right) = 0.633\mu$

The frequency of a photon may be calculated by using Equation 1 if its wavelength is known.

$$f = \frac{c}{\lambda} \tag{1}$$

where
- f = frequency.
- c = speed of light.
- λ = wavelength.

This equation is used in solving a problem in Example B.

EXAMPLE B: FREQUENCY OF A PHOTON

Given: Wavelength of a photon is 430 nm.

Find: Frequency of the photon.

Solution:

$$f = \frac{c}{\lambda}$$

$$= \frac{3 \times 10^8 \text{ m/s}}{4.3 \times 10^{-7} \text{ m}} \qquad (430 \text{ nm} = 4.3 \times 10^{-7} \text{ m.})$$

$$f = 6.98 \times 10^{14} \text{ Hz}$$

The energy of a photon may be calculated if its frequency is known by using Equation 2.

$$E = hf \tag{2}$$

where

E = energy of photon.
h = Planck's constant of 6.625×10^{-34} J x s.
f = frequency.

This equation is used to solve a problem in Example C.

EXAMPLE C: ENERGY OF A PHOTON

Given: Frequency of a proton is 5×10^{14} Hz.

Find: Energy of the photon.

Solution:

$$E = hf$$

$$= (6.625 \times 10^{-34} \text{ J x s}) \times (5 \times 10^{14} \text{ Hz})$$

$$E = 3.31 \times 10^{-19} \text{ J.}$$

EMISSION SPECTRA

Light in the optical spectrum can be produced by several means. The most important of these sources are electronic transistions within atoms, vibrational and rotational transitions within molecules, and incandescent emission from hot materials. Each of these produces a characteristic emission spectrum.

Figure 2 shows a typical line spectrum produced by electronic transitions in atoms. Intensity of light is plotted along the vertical axis, and the wavelength of light is plotted, in microns, along the horizontal axis. Line spectra are produced when an electric current flows through a low-pressure gas. Electron collisions raise atoms to high-energy states. Each atom then returns to lower states by emitting its excess energy in the form of photons. Emission lines in this spectrum vary greatly in width and intensity, but each atomic species has its own characteristic spectrum. Line spectra occur in the near IR, visible and UV portions of the electromagnetic spectrum.

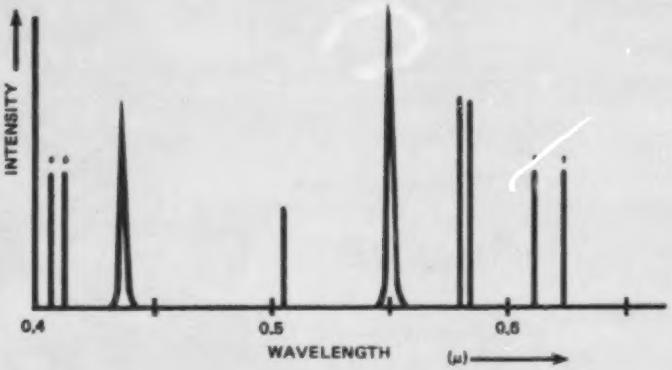


FIG. 2

Figure 3 displays a band spectrum produced by molecular transitions. Each band is actually composed of a large number of closely-spaced lines. Molecules contain energy not only in the electronic levels of their individual atoms, but also in the form of vibrational energy of the atoms and rotational energy of the entire molecule. Each band in Figure 3 is the result of a change in the vibrational energy of a molecule. Fine lines within the bands are produced by accompanying changes in rotational energy. Band spectra are produced by molecular gases such as diatomic nitrogen (N_2), by carbon dioxide (CO_2) and by some liquids and crystals. Band spectra occur throughout the optical spectrum but are most numerous in the infrared region.

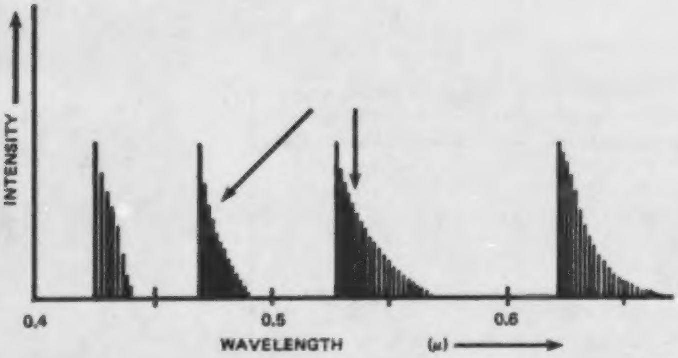


FIG. 3

Figure 4 shows a continuous spectrum emitted from a material at high temperature. Unlike line and band spectra, the continuous spectrum contains some light at all wavelengths. This light is the result of changes in the thermal energy of atoms in the material. In general, solids, liquids and high-pressure gases produce continuous spectra when heated to incandescence. Continuous spectra may be produced throughout the optical spectrum, with higher-temperature materials emitting shorter wavelengths. A more detailed discussion of this type of spectrum is contained in Module 13L3, "Light Sources."

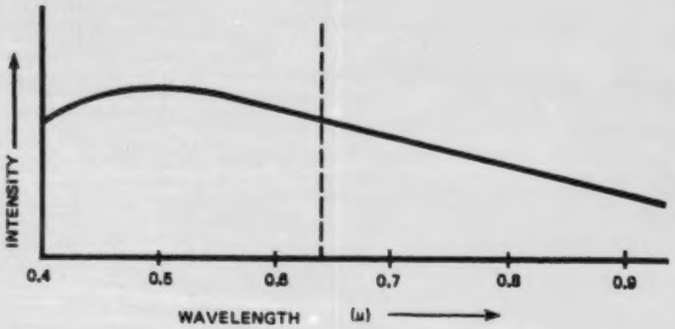


FIG. 4

APPLICATIONS OF THE OPTICAL SPECTRUM

The visible portion of the optical spectrum is the basis of human sight. All information received by the eye is carried by radiation in this portion of the spectrum. Fluorescent lights, television, photography and the candle are examples of technology applied to produce and use visible light.

The infrared portion of the spectrum cannot be detected by the eye. Instead, it is felt as heat radiation. Electrical heating devices, such as space heaters and toasters, produce heat in the form of infrared radiation. This spectral region contains much of the usable solar energy striking the earth. In recent years, a complex technology based upon infrared radiation has developed IR sensing and imaging devices used in night vision and temperature-measuring systems. Photographs taken with special IR-sensitive film are used to evaluate insulation in buildings and to locate hot spots in mechanical and electrical devices. Infrared technology is rapidly becoming a valuable tool in energy conservation and utilization.

Ultraviolet light, often referred to as black light, is also invisible. It is the component of sunlight that is responsible for suntans. Many materials absorb UV light and re-emit its energy in the form of visible light (longer wavelengths and less energy per photon). This process, called fluorescence, has many medical applications, and is used to locate leaks in fluid systems. Laundry detergents make clothes "whiter than white" because they leave a fluorescent residue that absorbs the UV in sunlight and re-emits the energy as visible light.

SUMMARY

The optical spectrum of light radiation includes both visible light and invisible light (ultraviolet and infrared). Light in the optical spectrum is produced by electronic transitions within atoms, by vibrational and rotational transitions within molecules and by incandescent emissions from hot materials. Each of these three sources produce, respectively, a characteristic emission spectrum that can be represented graphically as line, band or continuous.

LABORATORY

DEMONSTRATION OF DIFFRACTION GRATING

MATERIALS

Grating spectroscope.
Incandescent light bulb.
Spectrum tubes: helium, neon, carbon dioxide, water vapor.
Power supply for spectrum tubes.

PROCEDURE

The grating spectroscope is a device that separates light into its component colors and displays them against a calibrated scale for wavelength measurement. A detailed discussion of the principles involved is presented in Module 13L12, "Diffraction Grating - Measuring the Wavelengths of Light."

1. Turn on incandescent bulb. Place slit of spectroscope near bulb; look into eyepiece of spectroscope (Figure 5). Observe continuous visible spectrum; sketch and describe it in Data Table.

2. Replace incandescent source with helium spectrum tube and power supply (Figure 6). Caution: Power supply produces high voltage. Always turn off power supply before handling spectrum tubes. Also allow tubes to cool before removal, to avoid burn or breakage.

3. Turn on spectrum tube power supply. Position spectroscope with slit vertical near capillary portion of tube. Observe line spectrum of helium; sketch and describe it in Data Table.

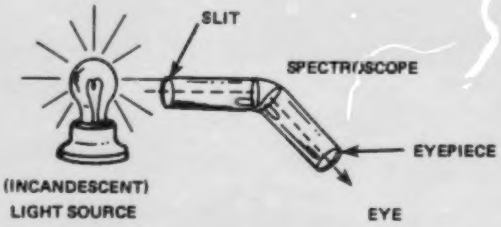
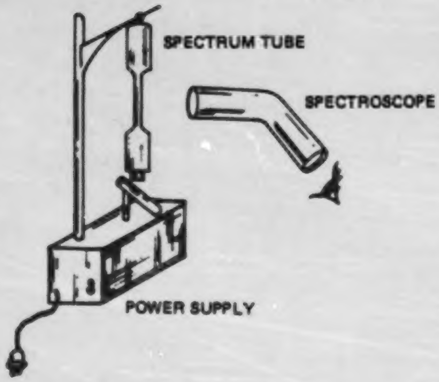


FIG. 5



4. Replace helim tube with neon tube, follow appropriate precautions. Observe, sketch and describe line spectrum.

5. Replace neon tube with carbon dioxide tube. Observe and describe band spectrum of carbon dioxide.

6. Replace carbon dioxide tube with water vapor tube. Observe, sketch and describe spectrum.

7. Replace helium tube in holder of power supply. Measure wavelengths of three strong lines; record in Data Table.

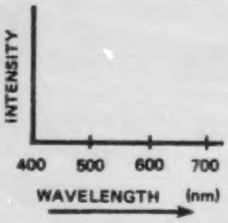
8. Use Equations 1 and 2 to calculate frequency and energy per photon of helium lines.

9. How do the sketches of each spectrum compare with those in the Discussion?

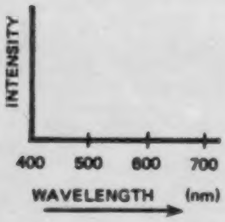
DATA TABLE

Description of Spectra:

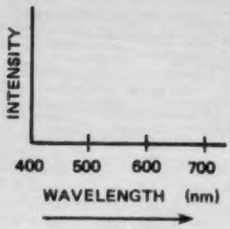
Incandescent lamp:



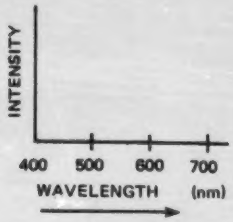
Helium:



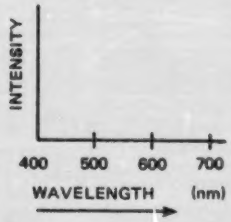
Neon:



Carbon dioxide:



Water vapor:

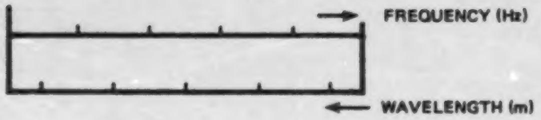


Helium lines:

Color	Wavelength (nm)	Frequency (Hz)	Energy of Photon (J)

SELF CHECK

1. On the following chart, label visible, IR and UV portions of the spectrum. Complete the wavelength and frequency scales by adding numerical values to the points given.



2. Define the following wavelength units:

   Micron.
   Nanometer.
   Angstrom.

3. Make the necessary calculations of energy, frequency and wavelength to complete the following chart.

Color	Energy of Photon (J)	Frequency (Hz)	Wavelength (nm)
Red			650
Green		5.78×10^{14}	
Blue	4.73×10^{-19}		

4. Explain briefly the origin of each of the following:

   Line spectra.
   Band spectra.
   Continuous spectra.

5. State an application of each of the following:

   Visible light.
   UV light.
   IR light.

6. Use the following experimental data to approximate the speed of light. The energy of a photon is found to be 3×10^{-20} J and its wavelength is independently measured to be 6.60×10^{-6} m.

7. Express the wavelength of a photon that has a frequency of 10^{14} Hz in Angstroms.

8. Write an equation giving the energy of a photon in terms of a constant divided by the wavelength of the photon.

9. Repeat problem 8 but now express the constant in units of eV·Å.

13L3

LIGHT SOURCES

Man has always needed light. Artificial production of light was one of the first technological accomplishments. Today it is essential to every part of modern civilization. This module will discuss the operation of several common devices for producing visible electromagnetic radiation. In the laboratory, the student will examine light from incandescent and fluorescent sources.

OBJECTIVES

Upon completion of this module, the student should be able to:

- Graph approximately the relative response of the human eye as a function of the wavelength of light.
- Describe the process by which light is produced and the output spectrum of the following light sources.

  Incandescent bulb.
  Fluorescent bulb.
  Metal vapor lamp.
- Draw three curves to show, on a single graph, the approximate "relative output" of a blackbody radiator at 3000° K, 4000° K and 5000° K, as functions of wavelength.
- Describe and give an example of a point source and an extended source of light.
- Given appropriate equipment, observe and describe the output spectrum and irradiance pattern of incandescent and fluorescent sources.

DISCUSSION

RESPONSE OF THE HUMAN EYE

Figure 1 shows the relative response of the human eye to light of various wavelengths. The eye is most sensitive to yellow light with a wavelength of 555 nm (nanometers). The response drops off rather sharply on each side of this value, reaching zero at 400 nm, the edge of the ultraviolet region, and at 700 nm, the edge of the infrared region. Visible light sources are devices that produce radiation in the region of the electromagnetic spectrum limited to the range of 400 nm - 700 nm.

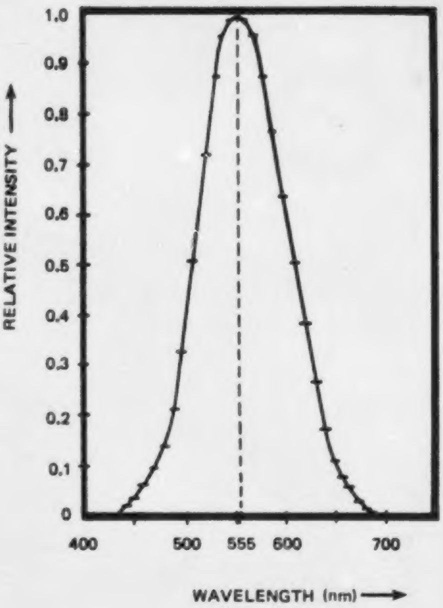


FIG. 1

EMISSION OF LIGHT BY COMMON SOURCES

Figure 2 illustrates an ordinary incandescent bulb. It consists of a thin filament of tungsten wire sealed inside a glass bulb envelope. The glass bulb is either evacuated or filled with an inert gas. When a current flows through the filament, it is heated to incandescence. Atoms in the filament are bound to fixed positions, but are free to vibrate around those positions. As this vibration (heat energy) occurs, neighboring atoms collide with one another violently. During these collisons, the atoms decelerate, and a portion of their vibrational energy is converted to electromagnetic radiation in the infrared, visible or near-ultraviolet portions of the spectrum. The result is a continuous spectrum (Figure 3).

Figure 3 shows the relative intensity produced by an incandescent source as a function of wavelength at three different temperatures. These curves are called blackbody radiation curves and are based upon the theoretical emission from a perfect radiator (blackbody). Although practical incandescent sources do not emit as much light as a perfect blackbody radiator would at the same temperature, the spectral distribution of the light produced is nearly the same. Notice that in Figure 3 an increase in temperature increases the intensity of light produced at all wavelengths. Note also that raising the temperature causes the peak of the emission curve to move toward the shorter wavelengths.

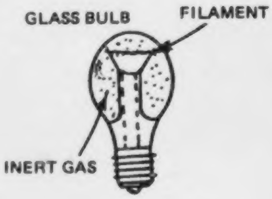


FIG. 2

The curves in Figure 3 also reveal the inefficiency of incandescent light sources. Only that fraction of the output that falls in the visible region is useful for human vision. Most of the energy used to heat an incandescent source results in production of unneeded ultraviolet and infrared

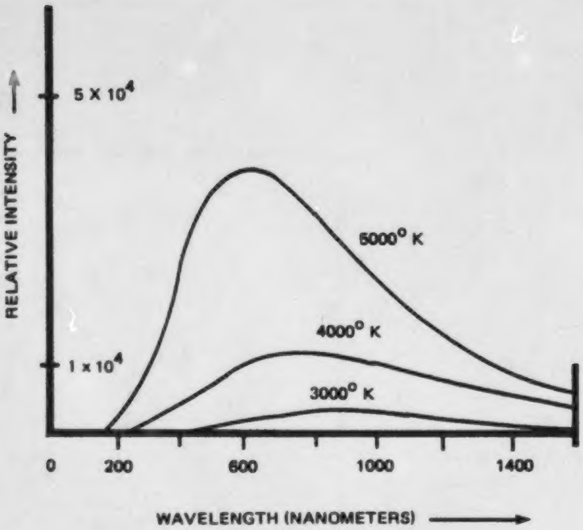


FIG. 3

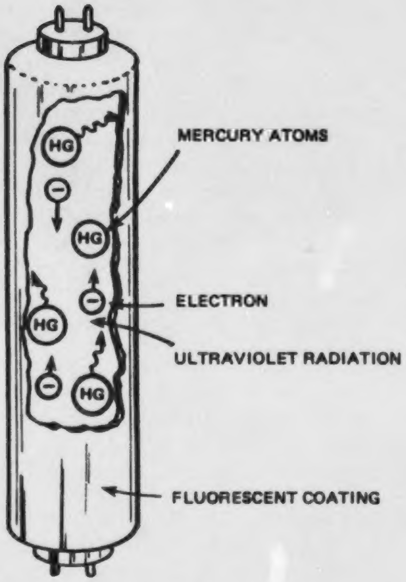

MERCURY ATOMS
HG
HG
HG
HG
ELECTRON
ULTRAVIOLET RADIATION
FLUORESCENT COATING

FIG. 4

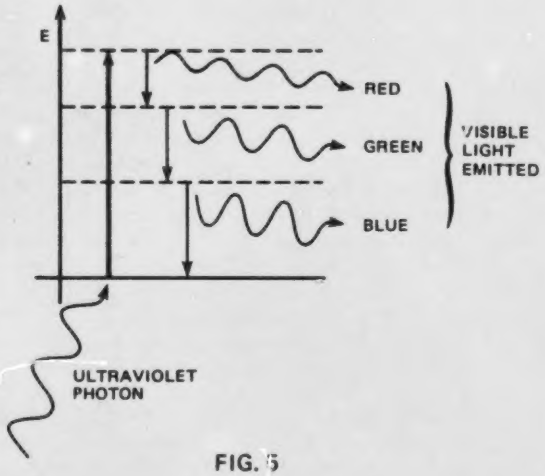

E
RED
GREEN
BLUE
VISIBLE LIGHT EMITTED
ULTRAVIOLET PHOTON

FIG. 5

radiation.

Figure 4 shows the operation of a fluorescent lamp. This lamp consists of a glass tube containing mercury vapor at low pressure. The inner surface of the tube is coated with fluorescent material. Alternating current is conducted through the tube by free electrons. When one of these electrons collides with a mercury atom, part of its energy is transferred to the atom, placing it in an excited energy state. The mercury atom usually emits this energy in the form of a photon in the ultraviolet range, although some visible light also may be produced. Ultraviolet photons strike atoms in the fluorescent coating of the tube and are absorbed, raising these atoms, in turn, to higher energy levels. These atoms release their energy in several steps, producing visible light, as shown schematically by the energy level diagram in Figure 5. The process is called fluorescence.

Several fluorescent substances may be used, and by proper selection and blending, can result in light of any desired color. Most fluorescent bulbs produce a continuous spectrum in the visible region. Their outputs also contain the strong visible lines of the mercury spectrum at 436 nm (blue) and 546 nm (green). Unlike incandescent sources, most of the energy required to activate fluorescent bulbs emerges as visible light, making them more efficient.

A third type of lighting, is the metal vapor lamp, which is important for large areas or for outdoor lighting. A number of metals that vaporize easily may be used, but mercury and sodium are more efficient than most others. In this lamp, a high pressure metal vapor is heated to incandescence. This produces the same line spectrum as an incandescent source. Mercury vapor lamps appear bluish because of the strong blue and green lines in the mercury spectrum. Sodium lamps are yellowish because of a pair of strong emission lines at 589 nm. Because this wavelength is near the peak of the response curve of the human eye, sodium vapor lamps produce usable light with greater efficiency than any other high-intensity source.

SIZE OF LIGHT SOURCES

Another important characteristic of any light source is its size. A point source of light is a source that emits all its radiant energy from a single point. Such an ideal source does not exist, although any source whose largest dimension is small compared with distance to the observer may be treated as a point source. Thus, a single incandescent bulb in a large room approximates a point source (Figure 6). This type of lighting produces well-defined shadows, and the light source may appear uncomfortably bright when viewed directly.

An extended source (Figure 7) is any light source whose dimensions are significiant when compared with the distance to

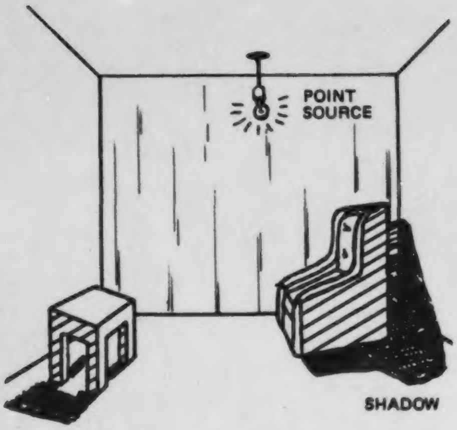


FIG. 6

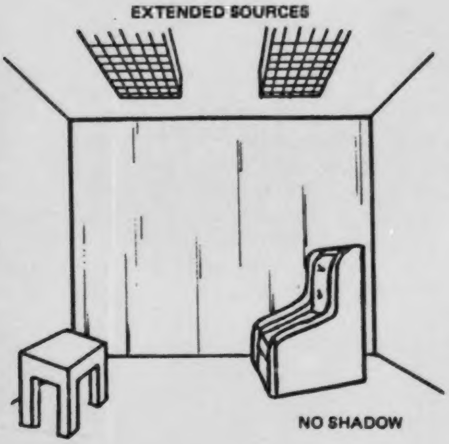


FIG. 7

the observer. Because light from an extended source emerges from a large area, the source does not appear uncomfortably bright when viewed directly. The shadows produced do not have well-defined boundaries, since light strikes objects from many directions. For these reasons, extended sources are preferred for most illumination applications. Lamp shades and home lighting fixtures for incandescent bulbs are examples of devices that change point sources into extended sources.

SUMMARY

Visible light sources produce electromagnetic radiation in the visible range 400 nm - 700 nm. Common emitters of visible light are the incandescent bulb (a bulb, tungsten filament, and inert gas), the fluorescent tube, (a fluorescent-coated tube containing mercury vapor) and the metal vapor lamp (a bulb of mercury or sodium vapor heated to incandescence). In size, a light source is either a point source that produces well-defined shadows, or an extended source over a large area with shadows of less contrast.

ISO-IRRADIANCE PLOT

MATERIALS

Clear incandescent bulb, 15 W or less, with socket and power cord.
Fluorescent bulb with support and electrical components (6 to 8 inch tube).
Two large sheets of white paper (approximately 2 x 3 feet).
Optical power meter.

PROCEDURE

1. Set incandescent bulb on one edge of a large sheet of paper (Figure 8). Mark position of point source.

2. Set power meter detector about 20 cm from light source and at same height as source. Measure irradiance (power/area); record reading. Mark position of detector.

3. Move power meter detector to new position at approximately same distance from bulb. Position radially to obtain same irradiance value. Mark position of detector.

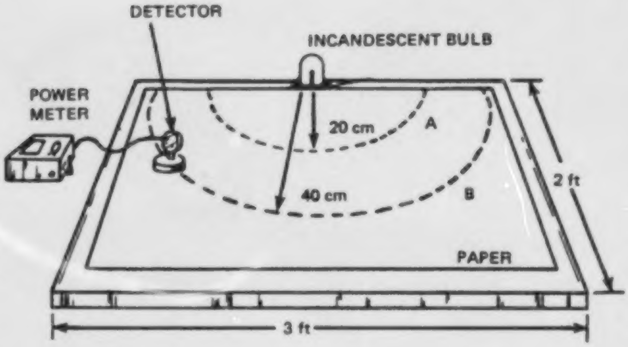


FIG. 8

4. Continue this process circumferentiatlly around light until ten points are marked. Connect all marks with a smooth line. (This line, line A in Figure 8, is called an iso-irradiance curve, since irradiance is the same at all points on it.)

5. Repeat steps 1-4 with a starting distance of 40 cm to draw curve B of Figure 8.

6. Refer to Figure 9. Set fluorescent bulb on one edge of a large sheet of paper. Mark position of bulb.

7. Repeat steps 1-6 above for fluorescent bulb. At each measurement point, power meter detector should be oriented to read maximum irradiance.

8. Sketch and describe in Data Table iso-irradiance plots of two light sources.

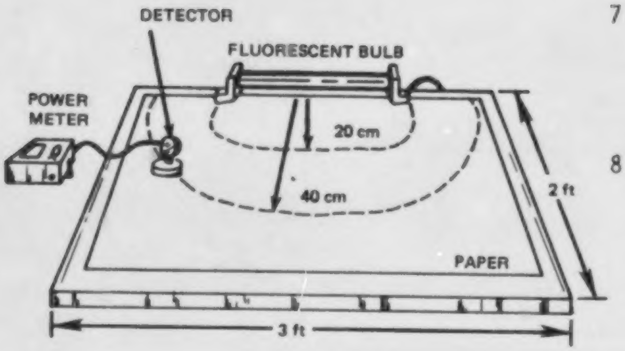

DETECTOR
FLUORESCENT BULB
POWER METER
20 cm
40 cm
2 ft
PAPER
3 ft

FIG. 9

SPECTRAL OUTPUT

MATERIALS

Clear incandescent bulb, 15 W or less, with socket and power cord.
Fluorescent bulb with support and electrical components.
Variable auto-transformer 0 to 125 V a.c.
Grating spectroscope.

PROCEDURE

1. Plug incandescent lamp into variable auto-transformer output. Set auto-transformer for maximum output. Set up spectroscope to view spectrum (Figure 10). Do not move spectroscope.

2. Turn off room lights. Set auto-transformer output to zero.

3. Observe spectrum of incandescent bulb as output voltage of auto-transformer is slowly increased. Note when various colors appear. Describe in Data Table.

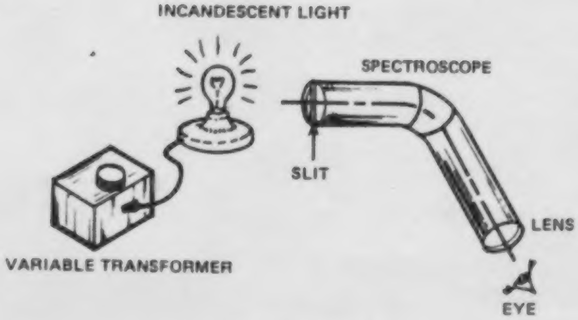

INCANDESCENT LIGHT
SPECTROSCOPE
SLIT
LENS
VARIABLE TRANSFORMER
EYE

FIG. 10

4. Replace incandescent lamp (and auto-transformer) with fluo-

cent lamp.

5. Turn on fluorescent bulb. Observe and describe spectrum.

DATA TABLE

1. Iso-irradiance plot.

Irradiance: 20 cm from incandescent bulb, E_1 = ____ mW/cm^2
40 cm from incandescent bulb, E_2 = ____ mW/cm^2
20 cm from fluorescent bulb, E_3 = ____ mW/cm^2
40 cm from fluorescent bulb, E_4 = ____ mW/cm^2

Sketch and describe iso-irradiance plots for:

Incandescent bulb. Fluorescent bulb.

2. Spectral output.

Describe how spectral output of incandescent lamp changed as voltage was increased.

Describe spectral output of fluorescent lamp.

SELF CHECK

1. List each of the following wavelengths of light:

   Shortest seen by human eye.
   Longest seen by human eye.
   Peak of human visual response curve.

2. Describe the process of producing visible light in each of the following:

   Incandescent bulb.
   Fluorescent bulb.
   Metal vapor lamp.

3. How does the output spectrum of a blackbody change as its temperature increases?

4. Identify each of the following as a point source or an extended source:

   Incandescent bulb viewed from five feet.
   Incandescent bulb viewed from fifty feet.
   Television screen in the same room.
   Star.

5. What is the relative response of the human eye to light of wavelength 650 nm?

6. The eye is most sensitive to light of wavelngth 550 nm. What is the frequency of this light?

7. By what factor does the intensity produced by the incandescent bulb of Figure 3 increase when the temperature is increased form 3000K to 5000K? (Use the values for 600 nm.)

8. The irradiance of a point source is inversely proportional to the square of the distance from the source. If the irradiance from a point source is 2×10^{-2} W/cm^2 at a distance of 30 cm from the source, what will the irradiance be 1 m from the source?

9. At what distance from the source of problem 8 will the irradiance be 1×10^{-4} W/cm^2?

LASERS

The laser is a source of unique light radiation, with a wide variety of applications. Laser light is used to drill holes in diamonds, to weld steel accurately, to measure large distances (for example, to the moon), to repair detached retinas and to accomplish many other applications. The laser is one of the more versatile tools of modern technology. This module will discuss the properties of laser light, the basic theory of operation of the laser and a number of important laser applications. In the laboratory the student will learn how to operate a low-power, helium-neon laser.

OBJECTIVES

Upon completion of this module, the student should be able to:

- Define the following properties of laser light:

  Monochromaticity.
  Directionality.
  Coherence.

- Describe the process of stimulated emission.
- State the purpose of the important elements of a laser.
- List several important applications of the laser.
- List all safety precautions necessary for safe operation of low-power, helium-neon laser; describe why each is important.
- Operate a low-power, helium-neon laser safely.

PROPERTIES OF LASER LIGHT

Light emitted by lasers is uniquely different from that produced by ordinary light souces, such as incandescent or fluorescent bulbs. All laser applications depend upon one or more of the unique properties of laser light.

Common light sources produce light that contains many different colors (different frequencies and wavelengths). Laser light is composed very nearly of only one wavelength. This property is called monochromaticity.

Ordinary light sources emit their light in all directions (Figure 1a). Light leaves a laser in a tight beam that spreads slowly as it moves away from the laser (Figure 1b). Thus, laser light possesses high directionality.

Individual waves of light produced by ordinary sources have no fixed phase relationships (Figure 2a). However, in the case of laser light, all the waves are in step. Coherence is the property of laser light, meaning that all light waves are in phase (Figure 2b).

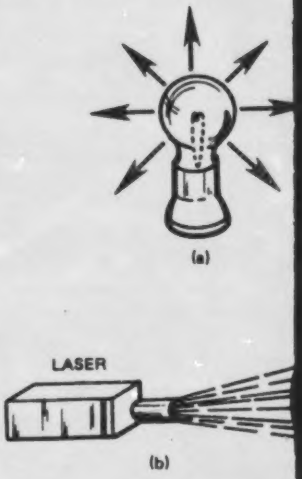


FIG. 1

STIMULATED EMISSION

The word "laser" is an acronym for Light Amplification by Stimulated Emission of Radiation. It is the process of stimulated emission that results in the unique properties of laser light.

Stimulated emission is shown schematically in Figure 3. Initially an atom is in a high-energy state E_3 represented by the solid circle in the figure. An incident photon, with an energy equal to the energy difference between this state and some lower-energy state E_2, passes near the atom. This stimulates the atom to drop to E_2, emitting the energy difference $(E_3 - E_2)$ in the form of a photon. This photon has the same wavelength as the stimulating photon (monochromaticity), travels in the same direction (directionality) and is in phase with the stimulating photon (coherence).

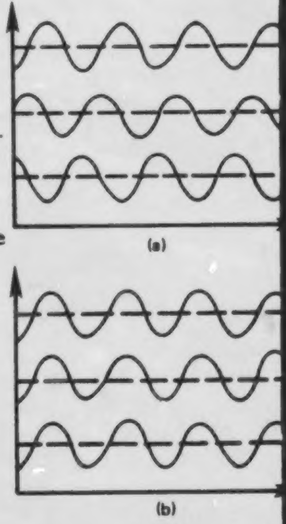


FIG. 2

ELEMENTS OF A LASER

For a laser to produce enough light to be useful, the stimulated emission process must occur many times. This repetition requires that a large number of atoms must be in energy state E_3 of Figure 3 and that a number of photons of the correct wavelength must be present for stimulation of these atoms. Figure 4 shows the four essential elements necessary to meet these requirements in any laser.

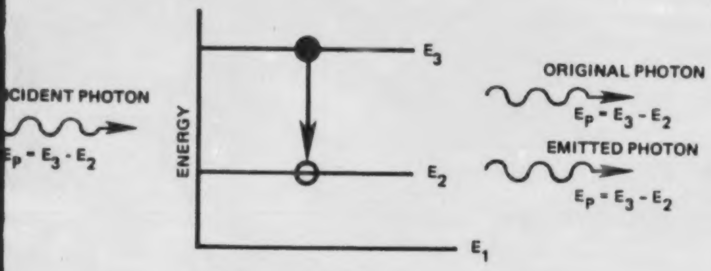

CIDENT PHOTON
$E_p = E_3 - E_2$
ENERGY
E_3
E_2
E_1
ORIGINAL PHOTON
$E_p = E_3 - E_2$
EMITTED PHOTON
$E_p = E_3 - E_2$

FIG. 3

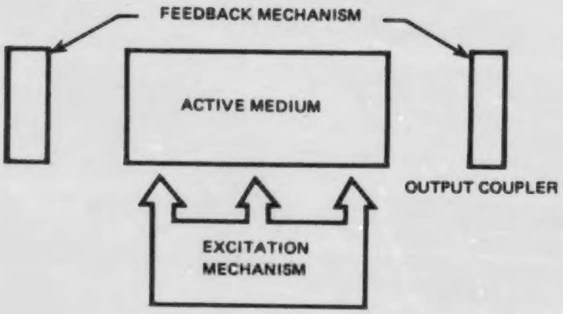

FEEDBACK MECHANISM
ACTIVE MEDIUM
OUTPUT COUPLER
EXCITATION MECHANISM

FIG. 4

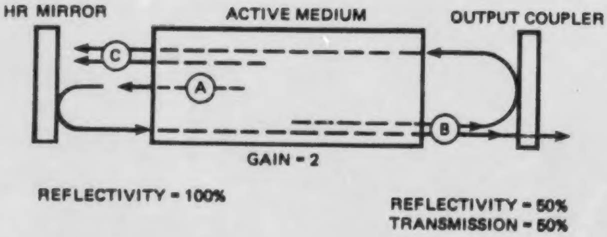

HR MIRROR
ACTIVE MEDIUM
OUTPUT COUPLER
C
A
B
GAIN = 2
REFLECTIVITY = 100%
REFLECTIVITY = 50%
TRANSMISSION = 50%

FIG. 5

The active medium is a collection of atoms or molecules that may be raised, or excited, to the proper energy level to produce light by stimulated emission. The excitation mechanism is an energy source that provides energy to the atoms, moving them to the proper energy level. This energy level is referred to as an excited state. The feedback mechanism is a pair of mirrors that reflect the coherent light produced back through the active medium to cause more stimulated emission. One of these mirrors is partially transparent and allows part of the light striking it to pass out of the laser. It is called the output coupler.

Figure 5 shows how the laser beam is produced. The output coupler of this laser is a partially-reflective mirror that reflects 50% of the laser light striking it and allows 50% to pass through to form the output beam. The other mirror is 100% reflective and is often called the high-reflectance (HR) mirror. The active medium acts as an optical amplifier with a gain of 2. The process begins at point A with the spontaneous emission of a photon traveling perpendicular to the mirror surfaces. This photon is reflected back through the active medium, where stimulated emission occurs. Two identical photons emerge in phase at point B. One passes through the output coupler to form the output beam. The other is reflected back through the active medium to produce more stimulated emission C. This process continues as long as there are enough atoms in the excited energy state to provide gain. The light produced is monochromatic, directional and coherent.

LASER CLASSIFICATION

While all lasers have the same basic elements and all operate as described above, the construction details and output characteristics vary greatly. Lasers may be classified in the following ways:

ACTIVE MEDIUM

Most common lasers use either a gas or a solid for an active medium. A gas laser (Figure 6) consists of a long, narrow tube that contains lasing gas. Mirrors are located at each end of the tube.

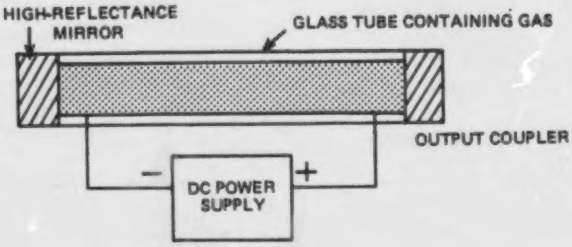

HIGH-REFLECTANCE MIRROR
GLASS TUBE CONTAINING GAS
OUTPUT COUPLER
−
+
DC POWER SUPPLY

FIG. 6

The active medium of a solid laser (Figure 7) consists of a cylinder of glass or of a crystal such as ruby. Mirrors may be located externally, as shown in Figure 7, or may be applied directly to the ends of the laser rod. Liquids and semiconductor materials may also be used as laser active mediums.

EXCITATION MECHANISM

Energy is supplied to most gas lasers by passing a current through the gas (Figure 6). A few gas lasers obtain energy from chemical reactions in the gas. Solid and liquid lasers obtain energy by absorbing light from another source (Figure 7), which is called optical pumping, and may employ either a flashlamp or a continuous lamp.

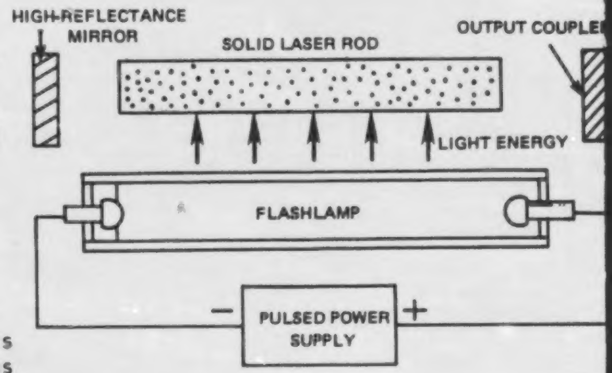

HIGH-REFLECTANCE MIRROR
SOLID LASER ROD
OUTPUT COUPLE
LIGHT ENERGY
FLASHLAMP
−
+
PULSED POWER SUPPLY

FIG. 7

OUTPUT DURATION

Continuous-wave lasers are those that produce continuous output beams. Pulsed lasers emit coherent light in short bursts. Lasers may be built to produce pulses with durations from less than a nanosecond to several milliseconds.

OUTPUT WAVELENGTH

Each laser emits a specific wavelength, but lasers may be built to produce light at almost any wavelength in the near-ultraviolet, visible and infrared portions of the spectrum.

LASER APPLICATIONS

There are many applications of lasers, and new ones are developed each year. A few of the more important applications are presented here:

- Alignment. The high directionality of light from lasers makes them ideal alignment tools. Low-power, helium-neon gas lasers are used for alignment problems throughout the industrial world, principally in construction industries.

- Materials processing. High-power, continuous, carbon dioxide gas lasers are commonly used for cutting, welding and heat treating of metals. Pulsed solid lasers are employed to drill small holes in metal parts. Lasers can drill small holes at a higher rate and with a smaller diameter than can be accomplished by any other means. A wide variety of materials, including cloth and semiconductor circuits, are cut with lasers.

- Communication. Modulated lasers are used as communication devices. A single laser beam can carry millions of telephone calls at once. Glass fibers carrying laser beams are replacing bulky, expensive cables used for telephone communication.

- Distance measurement. A laser rangefinder sends out a short pulse of light and measures the time necessary for a reflection to return. This technique is used in military applications and in surveying. Astronomical laser rangefinders have measured the distance to the moon and the altitude of satellites to within a few centimeters. Laser interferometers are used to measure short distances to within a millionth of an inch. This technique is discussed in Module 13L11, "Interference - Testing Flat Surfaces."

Other laser applications include eye surgery, controlled fusion research, three dimensional photography, laser weapons, light shows and cancer research. New applications are being developed continually.

LASER SAFETY PRECAUTIONS

Different lasers require different safety precautions. The following precautions must be observed in the operation of the low-power, helium-neon laser used in the laboratory portion of this module.

- Do not look directly into the laser beam. The lens of the eye will focus the laser beam to a small spot on the retina. Serious, permanent eye damage can result.
- Do not look at relfections of the beam from reflective surfaces. Mirrors, watch crystals, polished metal surfaces or other highly reflective surfaces reflect laser light with little decrease in the intensity of the beam. These reflections are as dangerous as the original beam.
- Take care when moving the laser or when moving objects into the laser beam. Direct and orient all beams and reflections to minimize the possibilities of accidental eye exposure.
- Beware of high voltages when operating the laser with case or enclosure opened. Laser power supplies contain high voltages. Electrical shocks account for more laser-related injuries than any other source.
- Operate the laser in a controlled area only. Be sure that everyone present is familiar with laser safety precautions.
- Do not track vehicles or aircraft with the laser beam. Such action could result in injury or property damage, and is forbidden by Federal law.
- Do not leave a laser unattended when it is operating or is in an operable condition. Always turn off the power switch and unplug the power cord. Store lasers where unqualified persons do not have access to them.

SUMMARY

Lasers produce concentrated coherent, light beams of a single wavelength. Beams are emitted when an excitation mechanism (d.c. power supply or light source) provides energy to the atoms of an active medium (gas, solid or liquid), causing photons to be reflected back and forth between mirrors. Some of this light emerges through one of the mirrors as a useful laser beam. Laser beams have either a constant or a pulsed output. Laser operators must observe strict safety precautions to avoid injury.

The laser is one of the more important inventions of the twentieth century. New and vital laser applications are being developed constantly.

OPERATION OF A LOW-POWER, HELIUM-NEON LASER

MATERIALS

Low-power, helium-neon laser, 2 mW or less.
Shutter.
Beam block (block of wood painted with flat black paint).
Allen wrench or screwdriver to remove laser cover.

PROCEDURE

1. Begin with laser power switch in off position, power cord unplugged and shutter closed.
2. Place beam block to intercept laser beam about 30 cm from laser output aperture (Figure 8).
3. Plug in laser power cord.
4. Turn on laser power switch; wait ten seconds.
5. Open shutter.
6. Observe laser beam. This can be done by lightly dusting chalk dust or blowing smoke into path of beam. (Check with instructor.)

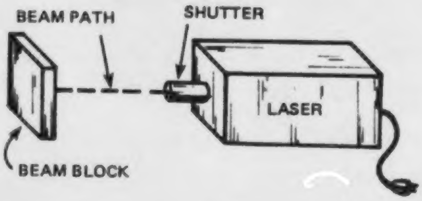


FIG. 8

7. Move beam block along beam path to a distance of about 2 m from laser. Observe divergence of beam.
8. Turn off power switch; close shutter; unplug laser.
9. Use allen wrench or screwdriver provided to remove screws from laser cover or end plate. Remove cover or slide laser from housing (depending on model).
10. Examine laser tube; sketch in Data Table.
11. Carefully reassemble laser. Test assembled laser by operating according to instructions above. Observe all safety precautions.
12. How is a laser beam produced by this gas laser? (Refer to

the Discussion and to Figure 6.)

DATA TABLE
Sketch laser tube. Label parts.

SELF CHECK

1. Match each of the following properties of laser light with the characteristic it describes:

(1) Monochromaticity.	a. All light waves are in phase.
(2) Directionality.	b. Light leaves the laser in a narrow cone and spreads very slowly.
(3) Coherence.	c. Laser light contains very nearly only one wavelgnth.

2. Draw and label a diagram representing the stimulated emission process.

3. Name each of the following elements of a laser:

   a. A source of energy to excite atoms to the proper energy level for stimulated emission to occur.
   b. A collection of atoms that can produce light by stimulated emission.
   c. A partially transparent mirror that allows part of the laser light to leave the laser.
   d. A pair of mirrors to reflect laser light to continue the stimulated emission process.

4. List several important applications of lasers.

5. List all safety precautions to be observed in the operation of a low-power, helium-neon laser.

6. What is the wavelength of orange laser light? (f = 5×10^{14} Hz)

7. A particular laser beam spreads 1 inch for every mile that it travels. If the beam is originally 0.01 inch in diameter what is the width of the beam 10 miles from the source?

8. A pulsed laser emits 3000 pulses per second. How long does it take for the laser to pulse 10 times?

9. A laser and a point light source both have a power rating of 1 watt. Why would it be dangerous to look directly into the laser beam but not to look directly at the point source?

10. How long would it take a laser beam to travel from New York to Los Angeles (approx. 3000 miles)?

13L5

STROBOSCOPES — LIGHT PULSES

Many applications of light radiation require that it be delivered in the form of short pulses. Photographic flashtubes, automobile timing lights and lasers produce bursts of radiant energy. This module discusses the stroboscope as a source of optical pulses. Both design features and applications of the stroboscope will be presented. In the laboratory, the student will use a stroboscope to measure the rotational rate of a fan and to observe water drop formation. He will also utilize an automobile timing light (strobe-light) to check the timing of an automobile engine.

OBJECTIVES

Upon completion of this module, the student should be able to:

- Explain how a stroboscope may be used to "freeze" repetitive motion or greatly reduce its apparent rate for observation.
- Diagram a simple strobscope, and state the purpose of each of the major parts.
- Explain how an automobile engine timing light functions.
- Given the appropriate equipment, measure the rotational rate of an electric fan with a stroboscope; and connect a timing light to an automobile engine and observe the timing mark.

548

DISCUSSION

The stroboscopic effect is based upon the human sense of vision. Sensors of the eye respond to short optical pulses, but the information is processed rather slowly. The human visual system cannot separate two optical pulses which reach the eye in a time interval of less than about one tenth of a second. A succession of optical pulses arriving at the eye at a rate of over ten per second is seen as an unbroken sequence, and any periods of darkness between these pulses cannot be detected by the observer. Television and motion pictures are two common applications of this principle. Both produce stationary images at a rate exceeding the resolving ability of human vision. Thus, images appear to move with an unbroken, continuous motion.

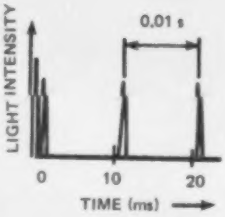


FIG. 1

The stroboscope is a device that produces short optical pulses at a predetermined rate. A typical stroboscope can be adjusted to emit from 1 to 100 pulses per second, with a pulse duration of about 5 microseconds. The time-varying output of a stroboscope at 100 pulses per second is shown in Figure 1. At this rate, the period of darkness between two adjacent pulses lasts two thousand times as long as one of the optical pulses.

Figure 2 shows how stroboscopic light can be used to "freeze" repetitive motion for observations and measurements. The object to be studied is a rotating disk upon which an arrow is painted. The stroboscope is set to produce one output pulse for each revolution of the disk. Figure 2a shows the disk illuminated by a brief pulse of light. Figures 2b and 2c show the disk continuing its rotation with no illumination. After one revolution is completed (Figure 2d), another pulse of light from the stroboscope illuminates the disk in its original position. This process continues, with an optical pulse being generated at the same point during each rotation of the disk. To a human observer, the disk appears to be stationary.

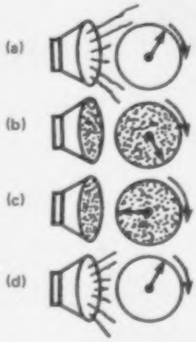


FIG. 2

As the pulse repetition rate and the rotation rate become lower, the time interval between pulses or rotation becomes greater. If the pulse repetition rate of the stroboscope is slightly less than the rotation rate of the disk (Figure 3), the disk will rotate slightly more than one complete revolution between light pulses. To an observer, the disk will appear to rotate in its actual direction of rotation at a greatly reduced rate. If the pulse repetition rate of the stroboscope is slightly greater than the rotation rate of the disk (Figure 4), the disk will be illuminated before it completes its revolution, and it will appear to rotate slowly in the direction opposite to its actual direction of rotation. This process can be used to make a detailed analysis of the repetitive motion of high-speed machinery during actual operating conditions.

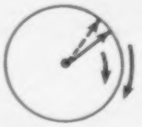


FIG. 3

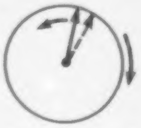


FIG. 4

A stroboscope may be used also as a light source for photography, to make a permanent record of motion. Figure 5 shows a photographic record of a falling body made with stroboscopic light.

The same techique is used to examine high-velocity motion, such as bullets in flight.

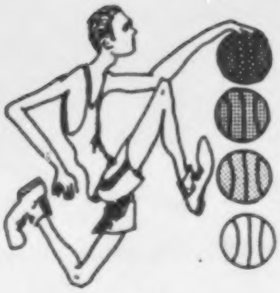


FIG. 5

Figure 6 shows a simplified diagram of a stroboscope. Light is produced by a current pulse through a flashlamp filled with neon or xenon. The internal resistance of this lamp provides the resistance of the RLC circuit. The d.c. power supply charges the capacitor to several hundred volts. This is not sufficient voltage to cause the lamp to conduct, but it does provide the short current pulse when the gas is ionized. Ionization is accomplished with a high-voltage pulse (several kilovolts) from a pulse circuit. This pulse is delivered to a trigger wire on the outside of the flashlamp. Timing of this pulse may be controlled by either an internal timer or some external trigger signal.

Figure 7 is a simplified diagram of an automobile-engine timing light, a common application of stroboscopic light. As in the more complex stroboscope, the d.c. power supply charges a capacitor to provide operating current for the lamp. The inductor is usually omitted to reduce circuit cost and size. Triggering is accomplished by connecting the high-voltage pulse from the number one spark plug directly to the trigger wire of the lamp. Thus, the lamp produces a pulse of light each time the spark plug fires, allowing precise timing for the firing of the plug.

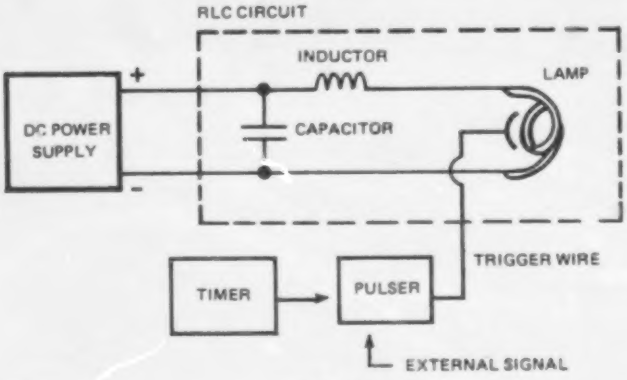


FIG. 6

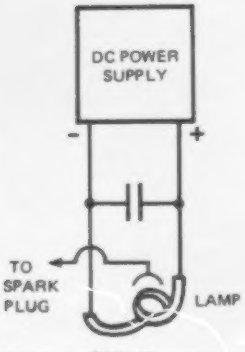


FIG. 7

SUMMARY

The stroboscope delivers light in short pulses useful in "freezing" action and in measuring the rate of action. Light is produced by a current pulse through a flashlamp filled with neon or xenon. A typical stroboscope is capable of emitting from 1 to 100 pulses per second, with a pulse duration of approximately 5 microseconds.

MEASURING ROTATIONAL RATE OF A FAN

MATERIALS

Stroboscope.
Electric fan.
Tape.

PROCEDURE

1. Place a strip of tape on one blade of electric fan.

2. Set stroboscope and fan as shown in Figure 8. Dim room lights.

3. Turn on fan and stroboscope. Adjust pulse repetition rate of stroboscope until fan blades appear to stand still. If tape strip appears on more than one blade, decrease pulse repetition rate until tape appears on one blade only (Figure 9). Read rotational rate (from stroboscope meter or from frequency-selection dial, depending upon stroboscope used); record this value in Data Table 1.

4. Set pulse repetition rate of stroboscope to twice that recorded above. Make fine adjustments to "stop" fan blades.

5. Continue above process, doubling pulse repetition rate each time. Explain appearances of the fan blade like the one shown in Figure 10. Record multiple frequences which just "stop" the blades.

6. True rotational rate of fan is the greatest pulse repetition rate for which tape strip appears on one blade only. Record true rotational rate in Data Table.

7. How could these laboratory principles be used to measure the rotational rate of a rotating piece of machinery?

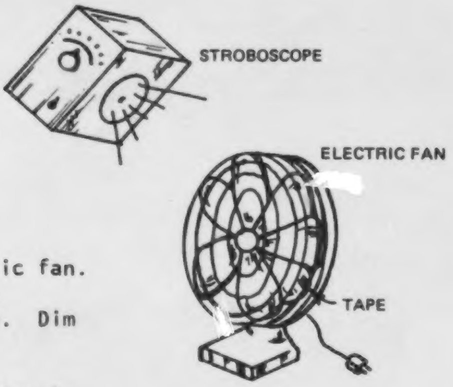


FIG. 8

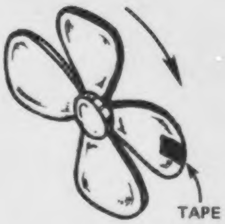


FIG. 9

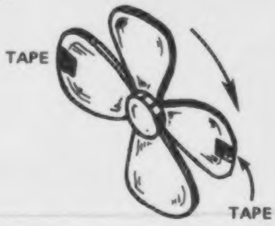


FIG. 10

OBSERVING DROP FORMATION

MATERIALS

Stroboscope.
Source of running water (tap).

PROCEDURE

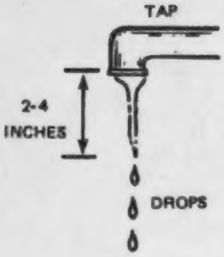


FIG. 11

1. Turn on water; obtain stream of drops (Figure 11). Unbroken stream of water should extend 2 to 4 inches from tap as shown.
2. Dim room lights; illuminate water stream with stroboscope.
3. Adjust pulse repetition rate of stroboscope until water drops appear to be stationary. Slight variations in frequency will cause drops to appear to either fall or rise slowly.
4. Observe drop formation.
5. How could photography with a strobe light "stop" the action of the falling drops?

OPERATING AUTOMOBILE ENGINE TIMING LIGHT

MATERIALS

Automobile.
Timing Light.

PROCEDURE

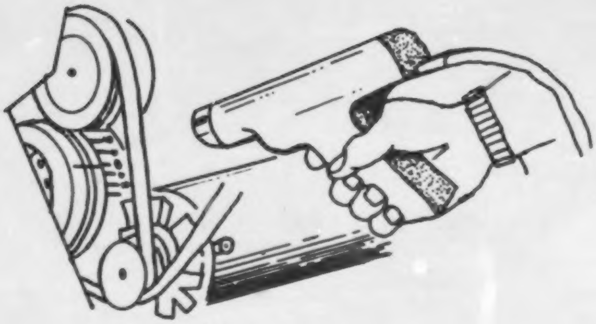
1. With automobile engine stopped, connect timing light as follows:

   Red wire to positive terminal of battery.
   Black wire to negative terminal of battery.
   Third wire to number one spark plug.

2. Be sure all wires are clear of moving parts; start engine.

3. Point timing light at timing marks at front of crankshaft (Figure 12). Press switch of timing light; observe alignment of moving mark (on crankshaft pulley or harmonic damper) with fixed marks.

4. How does this laboratory procedure demonstrate the principles in the Discussion of rotating disks and Figures 2, 3 and 4?

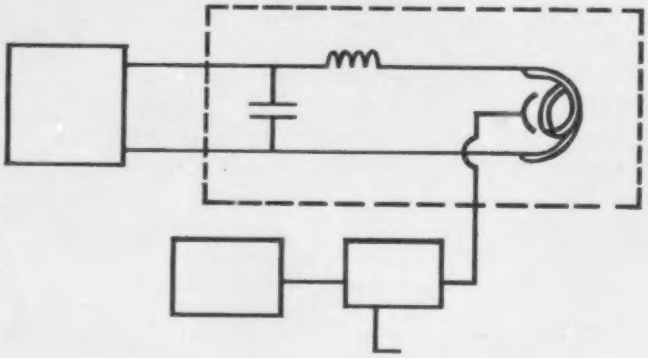


SELF CHECK

1. Indicate with a series of sketches showing a rotating body the timing of stroboscopic pulses of light to "freeze" the motion of the object.

2. Explain how a stroboscope may be used to observe high-speed repetitive motion at a greatly reduced rate.

3. Label the diagram on the right, showing the parts of a stroboscope



4. Draw and label a diagram, showing the parts of an automobile engine timing light.

5. A stroboscope light has a frequency of 120 Hz. What is the time between successive flashes. (Assume that the duration of the flash is negligible compared to the time between flashes.)

6. What is the maximum frequency of a stroboscope flash for which the human eye can distinguish the separate flashes?

7. A rotating disk has an arrow painted on it. (See Figure 2.) If the disk rotates at a frequency of 2000 rpm, what would an observer see if a stroboscope flashing at 1000 times per minute was used to illuminate the disk?

8. A stroboscope flashing 4000 times per minute is used to illuminate the rotating disk of problem 7. What would an observer see?

9. A stroboscope is set at a frequency of 7600 flashes per minute. What is the frequency of the stroboscope in Hz?

OPTICAL FILTERS — ABSORPTION OF LIGHT

Often the intensity or wavelength distribution of light produced by a light source does not meet the requirements of the application. In such cases, optical filters may be employed to absorb the unwanted electromagnetic radiation (light) while allowing the desired EM radiation to pass through. This module will discuss the absorption of light by matter and will present a description of several common types of optical filters. The exponential law of absorption will be explained and verified in the laboratory.

OBJECTIVES

Upon completion of this module, the student should be able to:

- Explain in a short paragraph how light is absorbed; include what happens to the energy after it is absorbed by the material.
- Explain with a diagram and in a short paragraph the exponential law of absorption.
- Given two of the following quantities, calculate the third;
  - Absorption coefficient of a filter at a specified wavelength.
  - Thickness of filter.
  - Transmission of filter at the specified wavelength.
- Describe in two or three sentences for each of the following filters how the transmission varies with changes in the wavelength of light:
  - Neutral density.
  - Cut-off.
  - Band-pass.
  - Color

- Given the optical density of a filter, calculate its transmission.

- Measure the transmission of an optical filter material as a function of thickness at a single wavelength and calculate its absorption coefficient given the appropriate equipment.

ABSORPTION OF LIGHT

Figure 1 shows a beam of light traveling through an optical filter. I_o is the intensity of the light incident upon the material and I is the intensity of the light emerging from the material. Equation 1 gives the transmission of the filter.

$$T = \frac{I}{I_o} \qquad (1)$$

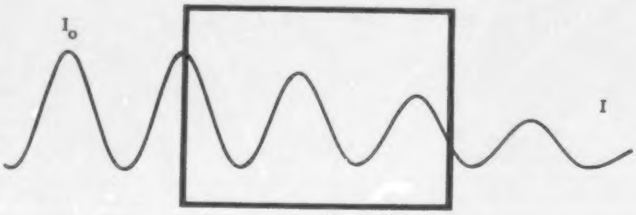


FIG. 1

where

T = transmission.
I_o = incident intensity.
I = transmitted intensity.

In some cases most of the light is transmitted (T = 0.999) and in others there is no transmission at all (T = 0), but the transmission is always less than one. The question that must be answered is: what happened to that portion of light energy that was not transmitted?

Figure 2 contains a simplified energy level diagram of an atom. The horizontal lines represent definite amounts of energy that the atom may contain. An upward arrow indicates an energy gain by the atom and a downward arrow indicates an energy loss. The atom is initially in the lowest energy state E_1, called the atomic ground state. The incident photon has an energy equal to the energy difference between energy levels E_1 and E_3. When the photon strikes the atom, it is absorbed, ceases to exist, and its energy appears as increased energy in the atom.

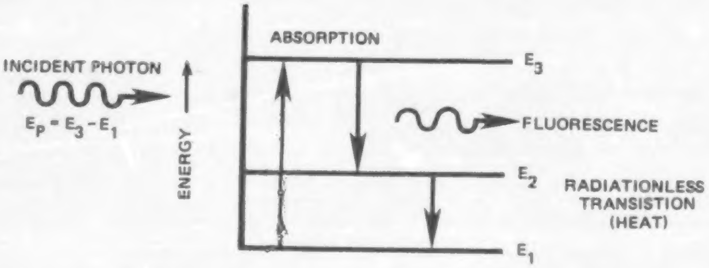


FIG. 2

Most atoms remain in an excited energy state for only a short time (average of 10^{-8} s).

They may release their excess energy by emitting it in the form of photons. This process is called fluorescence and is the mechanism used to produce visible light in fluorescent bulbs. Another energy release process is the radiationless transition shown in Figure 2. The electronic energy of the atom is converted to kinetic energy in liquids and gases or vibrational energy in solids. This appears as heat and causes a temperature rise in the material.

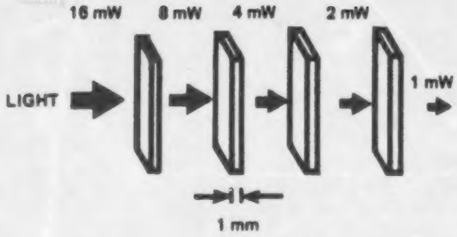


FIG. 3

EXPONENTIAL LAW OF ABSORPTION

It is obvious that increasing the thickness of an absorbing material will decrease the intensity of the transmitted light. Figure 3 shows light traveling through four pieces of identical filter material (each 1 mm in thickness). Each piece absorbs half of the light incident upon it. Figure 4 shows a graph of transmission versus thickness for this material. The transmission of an optical filter may be calculated using the exponential law of absorption given in Equation 2.

$$T = e^{-kx} \tag{2}$$

where

T = transmission.
k = absorption coefficient in cm^{-1}.
x = thickness in cm.
e = Naperian base (2.718).

In this equation the thickness x is always expressed in cm and the absorption coefficient k is in cm^{-1}. Example A illustrates how this equation is used in solving a typical problem.

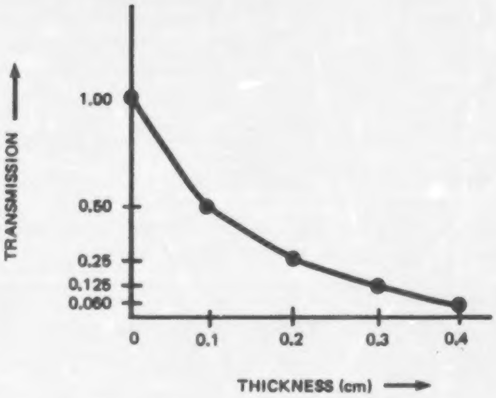


FIG. 4

EXAMPLE A: TRANSMISSION CALCULATION

Given: Absorption coefficient of the material in Figure 3 is $6.93\ \text{cm}^{-1}$.

Find: Transmission through a 2 mm thickness of this material.

Solution:

$$T = e^{-kx}$$

$$k = 6.93\ \text{cm}^{-1}$$

$$x = 2\ \text{mm} = 0.2\ \text{cm}$$

$$\therefore\ T = e^{-(6.93\ \text{cm}^{-1})(0.2\ \text{cm})}$$

$$= e^{-1.386}$$

$$T = 0.25$$

The absorption coefficient for a material may be calculated from Equation 3 if the transmission and thickness are known.

$$k = \frac{1}{x}\ \ell n\ \left(\frac{1}{T}\right) \qquad (3)$$

Example B shows how Equation 3 and Equation 1 are used in solving a problem.

EXAMPLE B: CALCULATING ABSORPTION COEFFICIENT

Given: A 10-mW laser beam strikes a piece of filter material 1.2 cm thick. 0.42 mW is transmitted.

Find: Absorption coefficient of material.

Solution:

$$T = \frac{I}{I_o}$$

$$= \frac{0.42\ \text{mW}}{10\ \text{mW}}$$

$$T = 0.042.$$

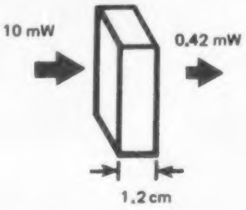
$$k = \frac{1}{x}\ \ell n\ \left(\frac{1}{T}\right)$$

$$= \frac{1}{1.2\ \text{cm}}\ \ell n\ \left(\frac{1}{0.042}\right)$$

$$= \frac{1}{1.2\ \text{cm}}\ \ell n\ (23.8)$$

$$= \frac{3.17}{1.2\ \text{cm}}$$

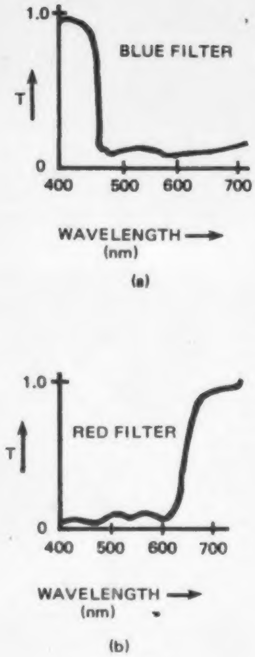
$$k = 2.64\ \text{cm}^{-1}$$



OPTICAL FILTERS

The absorption coefficient of any material (and thus its transmission) is a function of the wavelength of light striking the filter. Thus far, discussion in this module has dealt with thickness variations of absorbing materials at a single wavelength. This section assumes a fixed thickness and presents transmission variations as the wavelength of light changes.

Figure 5 shows transmission of two optical filters as functions of wavelength. The blue filter in Figure 5a allows most of the blue light to pass but absorbs other colors of the visible spectrum. The red filter (Figure 5b) allows only red light to pass through. These filters have a sharp division between high and low transmission regions and are called cut-off filters. They are used to eliminate unwanted wavelengths in many optical systems.



(a)

(b)

FIG. 5

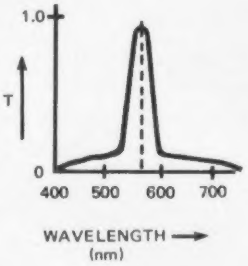


FIG. 6

Figure 6 is the transmission curve of a band-pass filter. This type of filter passes a narrow band of wavelengths but blocks all light outside this band. Darkroom safe lights use band-pass filters that pass only those wavelengths to which the film is not sensitive.

Transmission of a neutral density filter is shown in Figure 7. This filter is designed to have the same transmission for all wavelengths over a broad range of the spectrum. Neutral density filters can be made for any region of the optical spectrum. The gray lenses of most sunglasses are neutral density filters for visible light.

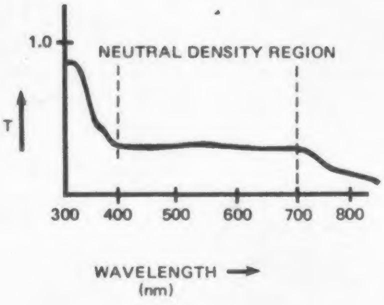


FIG. 7

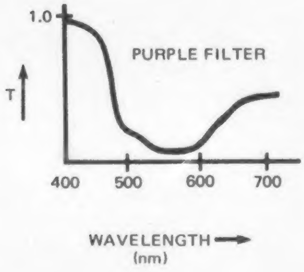


FIG. 8

Colored filters of glass or plastic are often used in lighting systems and decorative applications such as colored windows. Figure 8 is the transmission curve of a purple filter. The purple color is produced by a combination of blue and red light. Color filters are available in a wide variety of colors.

A term often used to describe the transmission of neutral density filters, laser safety goggles and other low-transmission filters is optical density. If the optical density of a filter is known, its transmission may be calculated from Equation 4.

$$T = 10^{-OD} \tag{4}$$

when
T = transmission.
OD = optical density.

Example C shows how this equation is used in solving a problem.

EXAMPLE C: CALCULATION OF TRANSMISSION FROM OPTICAL DENSITY

Given: The optical densities of two filters are:
a. 3.0.
b. 0.3.

Find: Transmission of each filter.

Solution:

a. $T = 10^{-3}$

$T = 0.001$

b. $T = 10^{-0.3}$

$T = 0.50$

SUMMARY

Whenever light passes through any material some of the light is absorbed. The fraction of the light that is transmitted through a material depends on the absorption coefficient of the material at the light wavelength and the thickness of the filter. Increasing the thickness of optical material reduces the transmission according to the Exponential Law of Absorption. Optical filters are constructed by selecting materials that have a low absorption coefficient for the wavelengths to be transmitted and a high absorption coefficient for the wavelengths that are to be blocked.

LABORATORY

MATERIALS

Low-power helium-neon laser, 2 mW or less.
Optical power meter.
Optical bench.
Filter holder.
Two carriages.
Lab jack.
Metric micrometer.
Plastic color filters of uniform thickness cut to fit filter holder (5 pieces of light red; 5 pieces of light green).

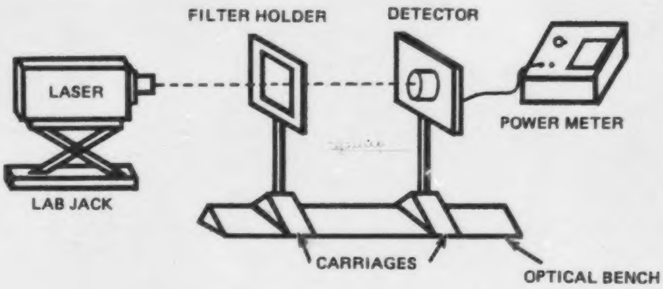
PROCEDURE

1. Review laser safety precautions.

2. Set up equipment as shown in Figure 9.

3. Observing all safety precautions, turn on laser; align beam to pass through center of empty filter holder and strike detector surface of power meter. Allow laser to operate for 15 minutes before taking data. This warm-up time will minimize output fluctuations.

4. Measure and record output power of laser with no filters in beam. If periodic variations occur in laser output, take power readings at maximum power attained. Use this procedure throughout the measurements.



5. Place one piece of red filter in filter holder. Measure and record transmitted power.

6. Place additional red filters in holder,

measuring power as each filter is added.

7. Remove red filters; repeat with green filters.

8. Turn off laser.

9. Use micrometer to measure thickness of one piece of filter of each color. Record thickness in data table.

10. Calculate transmission for each measurement of both red and green filters by dividing transmitted power by power recorded for laser with no filters.

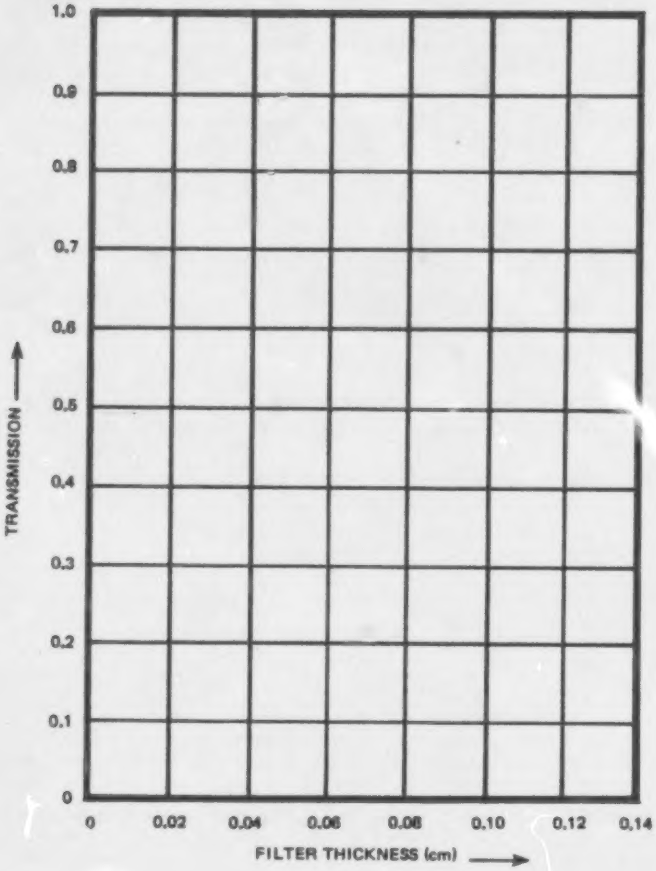
11. Calculate total filter thickness in each case by multiplying thickness of one filter by number of filters used.

12. Using Equation 3, thickness of five pieces of filter material and transmission of the five filters, calculate absorption coefficient of red and green filters for helium-neon laser light.

13. Draw curves showing transmission versus thickness for both filter materials on graph paper provided; label each curve and compare to Figure 4.

DATA TABLE

Red Filter Filter thickness = __________ cm.				Green Filter Filter thickness = __________ cm.			
Total Filter Thickness No. of Pieces	(cm)	Transmitted Power (mW)	Transmission	Total Filter Thickness No. of Pieces	(cm)	Transmitted Power (mW)	Transmission
0				0			
1				1			
2				2			
3				3			
4				4			
5				5			



Absorption coefficient. Red: k = _______ cm^{-1}

Green: k = _______ cm^{-1}

SELF CHECK

1. Draw and label a simple energy level diagram showing absorption of light by an atom. Include possible transitions that may remove the energy from the atom.

2. Draw a curve showing transmission of an optical filter as a function of thickness at one wavelength.

3. A filter 0.5 cm thick has an absorption coefficient of 6.3 cm^{-1} at a wavelength of 450 nm. Calculate its transmission at that wavelength.

4. A piece of glass 1.2 cm thick has a transmission of 0.82 at 633 nm. What is its absorption coefficient at that wavelength?

5. Draw a curve showing transmission of each of the following as a function of wavelength:

   Color filter.
   Neutral density filter.
   Cut-off filter.
   Band-pass filter.

6. The optical density of a filter is 2.0. What is its transmission?

7. For a certain material T = 0.26 for light with an intensity of 20 mW. What is the intensity of the transmitted light?

8. Use Equation 2 as a starting point in order to derive Equation 3.

13L7

SPECULAR AND DIFFUSE REFLECTION OF LIGHT

Most applications of light require that the direction of travel of the light waves be changed between the light source and some detector. The most common way to accomplish this is by reflection. The human eye itself is a complex detector system for analyzing light reflected from objects. This module presents the law of reflection of electromagnetic radiation and applications of that law to a variety of reflecting surfaces. Specular and diffuse reflectors and partial and total reflection will be discussed. Students will measure light reflected from specular and diffuse reflectors in the laboratory.

OBJECTIVES

To complete this module, the student should be able to:

- State the law of reflection and illustrate with a simple diagram.
- Define and illustrate specular reflection and diffuse reflection.
- Give three examples each of specular and diffuse reflectors.
- List the four processes that account for all light energy striking a surface.
- Explain in two or three sentences the purpose and operation of a beamsplitter.
- Given appropriate equipment, measure the distribution of laser light reflected from specular and diffuse reflecting surfaces.

DISCUSSION

The reflection of light may be studied by considering light to be composed of rays that travel in straight lines except for reflection or refraction at surfaces. Wave properties of light need not be considered when discussing reflection of light from most surfaces.

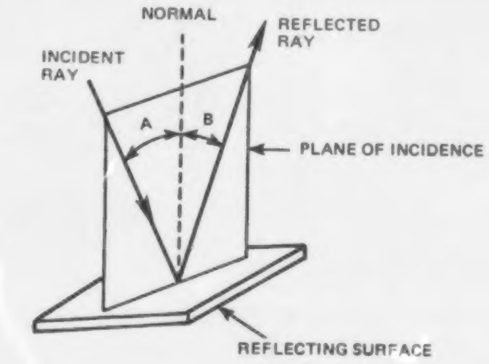
LAW OF REFLECTION

Figure 1 shows a light ray reflected from a plane surface. The normal line is a line perpendicular to the optical surface at the point of reflection. All angles are measured from this line. The incident ray, reflected ray and normal all lie in a plane called the plane of incidence. The law of reflection is given by Equation 1.

Law of reflection:

angle of reflection = angle of incidence,

or $\angle A = \angle B$ (1)



$\angle$A = ANGLE OF INCIDENCE
$\angle$B = ANGLE OF REFLECTION

FIG. 1

SPECULAR AND DIFFUSE REFLECTORS

A specular reflector is a smooth reflecting surface such as an ordinary mirror that reflects light according to the law of reflection and maintains internal order of the light beam. Figure 2 shows a light beam reflected from a plane specular reflector. In Figure 3 a similar beam is reflected from a curved specular reflector. In this case different parts of the beam are reflected at different angles but the internal order of the light rays is maintained after reflection.

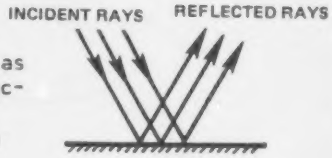


FIG. 2

A diffuse reflector is a rough reflecting surface, such as this sheet of paper, that reflects light randomly in all directions and does not maintain internal order in a light beam. Figure 4 shows reflection of light from such a surface. Each flat surface element of a diffuse reflector acts as a tiny mirror, and reflections from each element obey the law of reflection. Because of the random orientation of the flat surface elements, diffuse reflecting surfaces reflect light in all

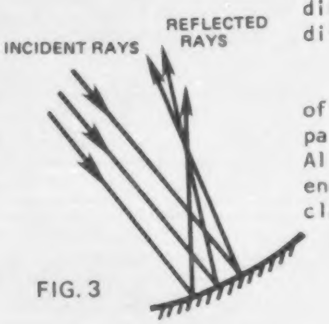

INCIDENT RAYS
REFLECTED RAYS

FIG. 3

directions. Figure 5 shows the radiation pattern from a diffuse reflector when it is struck by a beam of light.

Many surfaces display a combination of the characteristics of specular and diffuse reflectors. Figure 6 shows the typical pattern of light reflected from a glossy painted surface. Although some diffuse reflection takes place, most of the light energy is found in specular reflection and this surface is classed as a specular reflector.

PARTIAL AND TOTAL REFLECTORS

When light strikes a surface, some of the light energy is always reflected. But no surface can reflect all of the light that strikes it. Figure 7 shows light striking an optical surface. Part of the light is reflected (1), part is transmitted (2), a third portion is scattered due to diffuse reflections from surface imperfections (3), and a fourth part is absorbed by the optical material. The total energy of the incident light beam must be accounted for in these four processes as given in Equation 2.

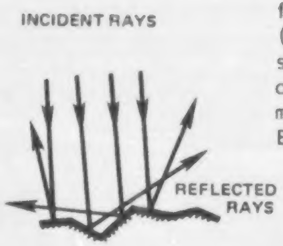

INCIDENT RAYS
REFLECTED RAYS

FIG. 4

Incident energy = scattered energy + transmitted energy + absorbed energy + reflected energy

or

$$E_I = E_S + E_T + E_A + E_R$$

or

$$E_I = S + T + A + R. \qquad (2)$$

Example A shows the use of this equation in solving a problem.

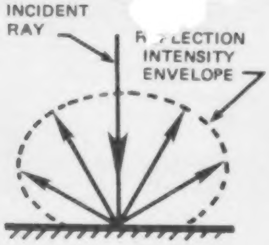

INCIDENT RAY
REFLECTION INTENSITY ENVELOPE

FIG. 5

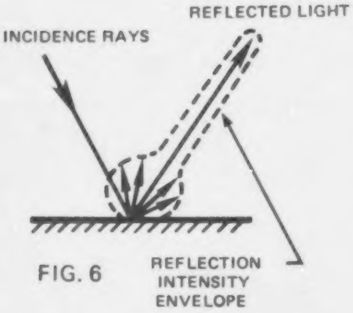

INCIDENCE RAYS
REFLECTED LIGHT
REFLECTION INTENSITY ENVELOPE

FIG. 6

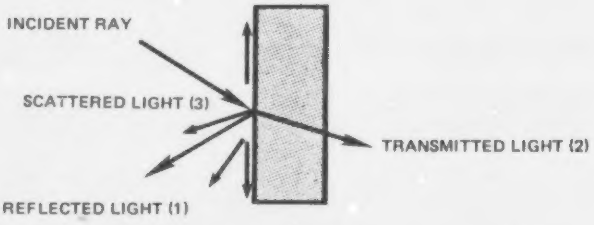

INCIDENT RAY
SCATTERED LIGHT (3)
TRANSMITTED LIGHT (2)
REFLECTED LIGHT (1)

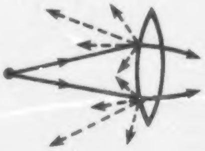
FIG. 7

EXAMPLE A: REFLECTION LOSS OF A LENS

Given: A lens has a transmission of 92%. The total of absorption and scattering is 0.5%.

Find: Total reflection loss.

Solution:



$$E_I = S + T + A + R$$

$$100\% = 0.5\% + 92\% + R$$

$$R = 100\% - 92.5\%$$

$$R = 7.5\%$$

Some light energy always goes into each of the four processes. Scattered light is undesirable in most optical systems and can be reduced to near zero intensity (0.01%) by polishing the optical surfaces and keeping them clean. Clean glass surfaces scatter so little light that scattering usually can be ignored. Absorption also reduces the intensity of the transmitted beam since it removes light energy from the incident beam Ordinary bathroom mirrors absorb 10 to 15% of the light that strikes them. In many highly transmitting optical components, absorption is reduced to almost zero. Ordinary glass windows absorb very little light. In windows maximum transmission is desirable while in other applications no light is transmitted.

A portion of the light is always reflected. Ordinary glass reflects about 4% of the light that strikes it at each of its surfaces. Figure 8 shows light striking a rear surface mirror. About 4% of the light is reflected by the glass surface at point A. The main reflected beam leaves the glass at point B, with approximately 92% of the original beam.

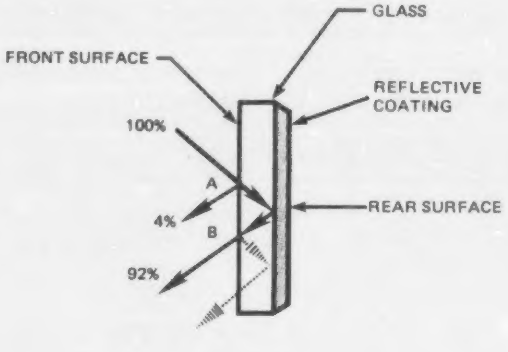


FIG. 8

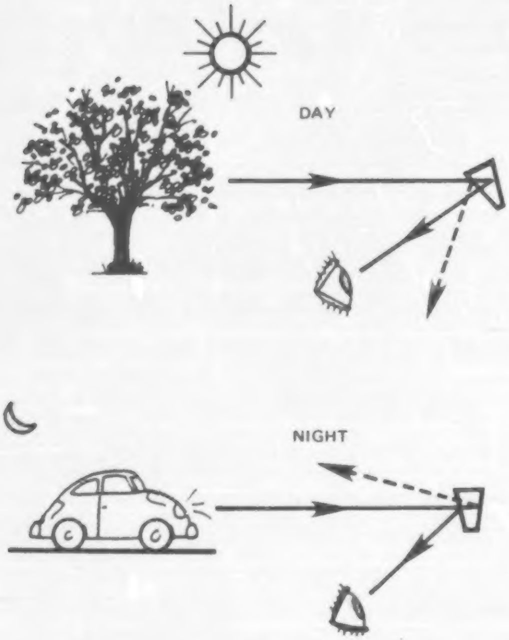


FIG. 9

About 3.8% of the original beam is reflected back into the glass at point B. Most of this light passes out of the glass after the next reflection. The two weaker reflected rays often form ghost images in mirrors. This can be eliminated by applying the reflective coating to the front surface of the mirror, but this also exposes the reflective coating to possible damage.

Figure 9 shows how partial reflection is employed in rear vision mirrors for automobiles. Such mirrors are actually optical wedges with the rear surface silvered. During the day the more intense reflection from the silver coating is viewed. At night the weaker reflection from the front surface is used to reduce light intensities from the headlights of trailing cars.

Coatings are often applied to optical surfaces to change their reflectivity. Highly reflective laser mirrors may reflect as much as 99.95% of the laser light that strikes them. Camera and telescope lenses often have antireflection coatings to reduce reflections to less than 1%.

BEAMSPLITTERS

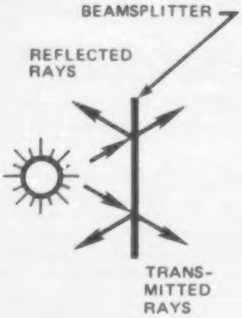


FIG. 10

In many applications it is desirable to reflect only a portion of the light striking a surface while allowing the remainder to pass through. This is accomplished with the partially reflective optical surface of a beamsplitter shown in Figure 10. The ratio of transmitted to reflected light may be set at any desired value. Beamsplitters are used in many optical instruments (including interferometers and range finders) but also have several more common applications. The partially reflective windows of many office buildings are in actuality applications of beamsplitters (Figure 10).

Figure 11 shows an electronic game using a beamsplitter. The gun is a light source that sends a beam toward a target. Part of the beam reaches the target in the back of the cabinet, but the score is recorded when the reflected beam strikes a detector in the base of the cabinet. The operator sees only the target and is unaware of the beamsplitter or detector.

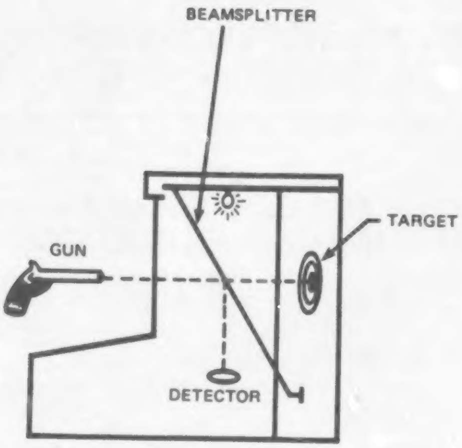


FIG. 11

SUMMARY

Whenever light strikes a surface it may be reflected, scattered, absorbed or transmitted. In most common optical components absorption and scattering account for only a small fraction of the total light energy. Thus, most of the light is reflected from the surface or transmitted through the optical component.

A diffuse reflector is a rough surface that reflects incident light in all directions. A specular reflector is a smooth mirror-like surface that reflects light rays according to the law of reflection:

angle of incidence = angle of reflection

Some light is always reflected from any optical surface. Mirrors reflect most of the light while beamsplitters reflect a portion and allow the remainder to be transmitted.

SPECULAR AND DIFFUSE REFLECTIONS OF A LASER BEAM

MATERIALS

Low-power helium-neon laser (2 mW or less).
Optical power meter (Metrologic #60-530 or equivalent).
Protractor.
Large compass (7-inch minimum), or string and tack.
Ruler.
Plane front surface mirror (Edmund #40.040 or equivalent).
White index card glued to block of wood.
White enamel surface (flat metal painted with gloss enamel paint).
Sheet of paper 8 1/2 inches x 14 inches.

PROCEDURE

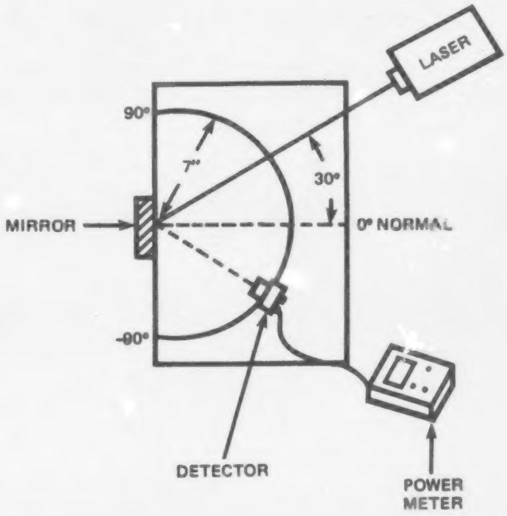
CAUTION: AVOID LOOKING INTO LASER BEAM, EITHER DIRECTLY OR WHEN BEAM IS REFLECTED OFF A SURFACE. THE BEAM CAN CAUSE EYE DAMAGE.

1. Use compass to draw a semicircle with a radius of about 7 inches on sheet of paper as shown in Figure 12. From center of semicircle, draw radial lines every 10° to correspond to graph paper in graph section of data.

2. Arrange equipment as shown in Figure 12. Observe all laser safety precautions. Laser beam should be level and should strike mirror surface at center with angle of incidence of 30°.

3. Dim room lights.

4. With laser and mirror stationary move power meter detector; measure reflected optical power at 10° intervals along semi-



circle. Record power in data table. Be sure detector surface is always at same distance from point of reflection and perpendicular to reflected light. Some measurements will require most sensitive scales of power meter. Be careful not to drive needle past full scale.

5. Replace mirror with index card; repeat entire procedure for diffuse reflection from paper.

6. Replace index card with enamel surface; repeat. Record data in combination column in data table.

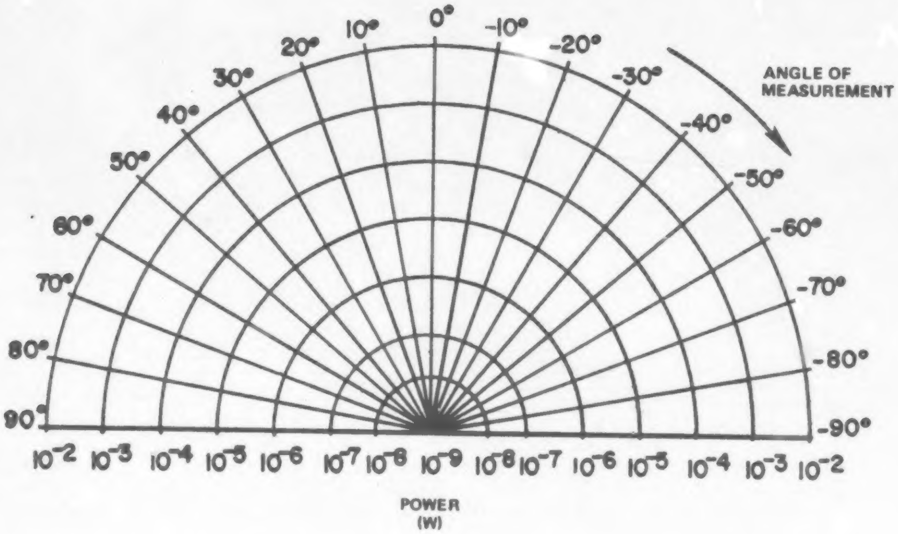
7. Plot data on graph paper provided in graph section. Plot power recorded along radial line at approximate angles. Radial distance on graph is therefore power in watts measured on the logarithmic scale of the power meter. Connect all data points with smooth curve.

8. Compare the resulting intensity patterns to those predicted in the discussion. (See Figures 5 and 6).

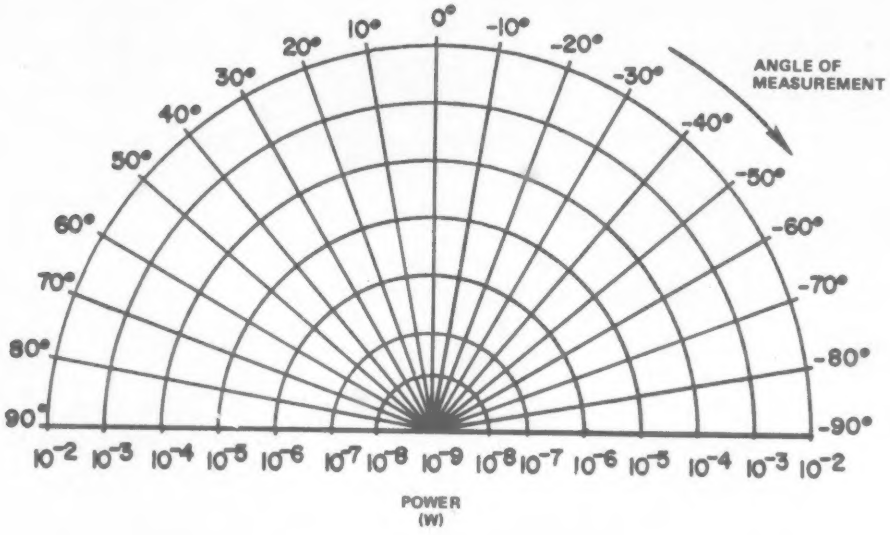
DATA TABLE

Angle of Measurement (degrees)	Reflected Power Measured (W)		
	Specular	Diffuse	Combination
-90			
-80			
-70			
-60			
-50			
-40			
-30			
-20			
-10			
0			
10			
20			
30			
40			
50			
60			
70			
80			
90			

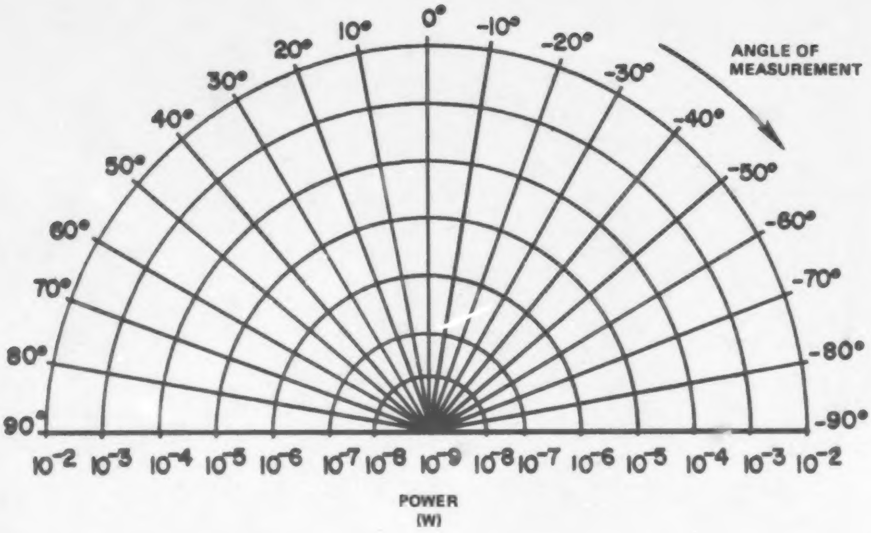
Specular Reflector



Diffuse Reflector



Third Reflector (Combination)


0°
10°
20°
30°
40°
50°
60°
70°
80°
90°
-10°
-20°
-30°
-40°
-50°
-60°
-70°
-80°
-90°
ANGLE OF MEASUREMENT
10^{-2} 10^{-3} 10^{-4} 10^{-5} 10^{-6} 10^{-7} 10^{-8} 10^{-9} 10^{-8} 10^{-7} 10^{-6} 10^{-5} 10^{-4} 10^{-3} 10^{-2}
POWER
(W)

SELF CHECK

1. Draw and label a diagram illustrating the law of reflection.

2. Draw and label diagrams showing light reflected from specular and diffuse reflecting surfaces.

3. Identify the following as specular or diffuse reflectors:

   Bathroom mirror.
   Sheet of paper.
   Ceiling tile.
   Watch crystal.
   Shiny linoleum tile.

4. Draw and label a diagram of light striking a beamsplitter at an incident angle of about 30°. Show both the transmitted and the reflected ray.

5. A beamsplitter reflects 49% of the light striking it and transmits 48% of the light. What happens to the other 3%?

6. Name the four processes that account for all light striking a surface.

7. What is the transmission of a lens that scatters 2%, reflects 10%, and absorbs 1% of the incident radiation?

8. If the incident intensity on the lens of problem 7 is 40 mW what is the intensity of the reflected light?

9. Light is incident upon a lens at an angle of 37° above the normal to the lens. At what angle should a collector designed to collect the reflected radiation be placed?

10. What is the intensity of light incident upon a lens if the intensity of the transmitted light is 27 mW, and the combined intensity of the absorbed, reflected, and scattered light is 4 mW.

11. What is the maximum percent of incident light that can be transmitted if 3% of the incident light is absorbed?

13L8

IMAGE FORMATION WITH MIRRORS

Mirrors are devices that change the direction of travel of light rays by reflection. Mirror surfaces are usually plane, spherically convex or spherically concave. Such mirrors may be used to focus light to a small area, collimate light beams, form images, or simply change the direction of a beam of light. This module discusses reflections from plane and shperical surfaces, as well as the images formed by these reflections under different conditions. In the laboratory, the student will use a concave mirror to construct a solar furnace and to produce real and virtual images.

OBJECTIVES

Upon completion of this module, the student should be able to:

- Define:
    - Real image.
    - Virtual image.
    - Focal point of a concave mirror.
    - Focal point of a convex mirror.
- Draw and label ray diagrams that illustrate:
    - Virtual image formed by a plane mirror.
    - Virtual image formed by a convex mirror.
    - Virtual image formed by a concave mirror.
    - Real image formed by a concave mirror.
- Given two of the following quantities, calculate the third:
    - Focal length of mirror.
    - Object distance.
    - Image distance.

- Given three of the following quantities, calculate the fourth:

  Object distance.
  Image distance.
  Object height.
  Image height.

- Given appropriate equipment, construct a solar furnace.

- Given the appropriate equipment, demonstrate formation of real and virtual images formed by a concave mirror.

MIRROR SHAPES

Although mirrors may be constructed in a wide variety of configurations, the three basic surface shapes most commonly used are used are plane, spherical concave, and spherical convex.

Plane mirrors (Figure 1) have plane reflecting surfaces. When incident parallel light rays strike a plane mirror, angle of incidence A and angle of reflection B are equal, and the same for all rays. Thus, the reflected rays leave the mirror parallel to one another as shown.

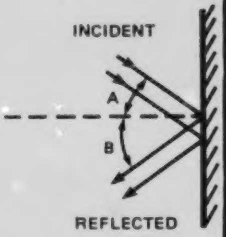


FIG. 1

Figure 2 shows a concave mirror. The optical axis, a line perpendicular to the mirror surface, usually is located at the center of the mirror. Point C is the center of curvature of the mirror surface: distance r is the radius of curvature of the surface.

When light rays parallel to the optical axis strike such a mirror, the equal angles of incidence and reflection are greater near the edge of the mirror than near the center. The reflected rays converge to a point F, called the focal point of the concave mirror. Distance f, from point F to the mirror surface, measured along the optical axis, is called the focal length of the mirror. This focal length is always one half the radius of curvature of the mirror surface (for spherical mirrors).

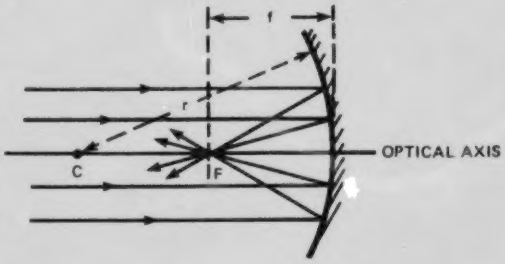


FIG. 2

Figure 3 shows a convex mirror. The optical axis and center of curvature are defined as for a concave mirror. Also, as with the concave mirror, parallel light rays have greater angles of incidence and reflection near the mirror edges. The focal point of a convex mirror is that point from which all reflected rays appear to have originated when parallel rays are incident upon the surface (located by dotted lines behind the mirror). As with the concave mirror, the radius of curvature is twice the focal length.

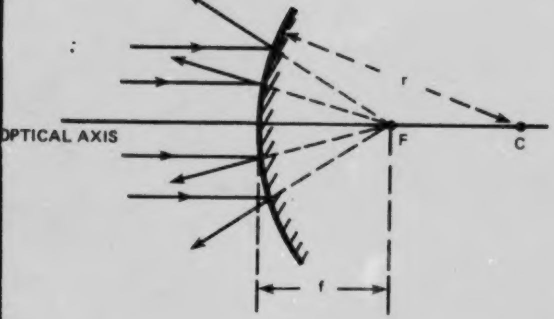


FIG. 3

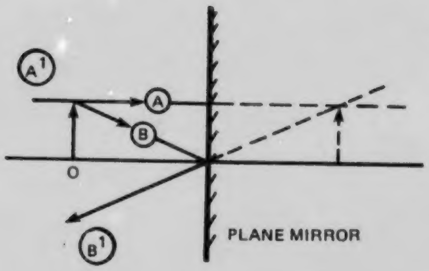


FIG. 4

IMAGE FORMATION

Figure 4 shows the formation of an optical image by a plane mirror. Object O is an arrow with its base on the optical axis. Two light rays are shown leaving the top of the object. Ray A, which is perpendicular to the mirror surface, is reflected back along itself to the top of the object (see A'). Ray B, which strikes the mirror surface on the optical axis, is reflected according to the law of reflection, with the angle of incidence equal to the angle of reflection. The two reflected rays A' and B' diverge as they leave the mirror. If their lines of travel are extended backward past the reflecting surface (dotted lines behind the mirror), they meet to form the top of the image I. The bottom of the image is located on the optical axis.

The image formed behind the reflecting surface is called a virtual image. It may be viewed directly from the front of the mirror, but cannot be focused onto a screen because there is actually no light behind the mirror. This type of image is the only type that can be formed with a plane mirror.

Figure 5 shows the formation of a different type of image by a concave mirror. The three rays used in this figure may always be used to locate images formed by spherical mirrors. They have the following characteristics:

Ray A is the light ray from the top of the object, through the center of curvature of the mirror, to the mirror surface. It will always be perpendicular to the mirror surface and is reflected back along itself.

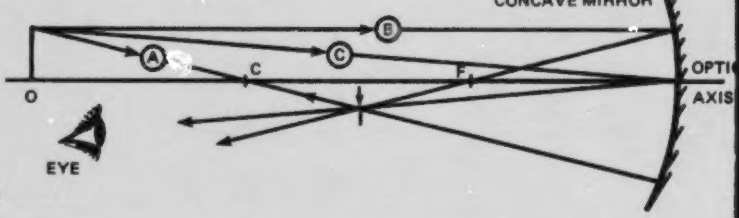


FIG. 5

Ray B, which is parallel to the optical axis, connects the top of the object to a point on the mirror surface. This ray is reflected through focal point F of the concave mirror.

Ray C connects the top of the object to the point of intersection of the optical axis and the mirror surface. Like all rays reflecting from the surface, it is reflected according to the law of reflection.

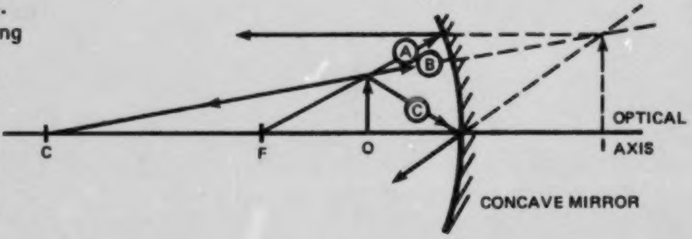


FIG. 6

The point at which all three rays intersect is the top of the image. It is possible to locate the image accurately with only two rays, but three usually are shown. The third may be used as a check.

The image in Figure 5 is called a real image. It is located in front of the mirror surface and, since it is formed by real rays of light, it may be focused on a screen placed at point I. The same real image may also be viewed directly from a point on the optical axis at a greater distance from the mirror than the image itself. (See "eye" in Figure 5.) The light rays entering the eye will diverge from I just as they would if it were a real object.

Figure 6 shows formation of a virtual image when the object is located inside the focal point of a concave mirror. In this instance, ray A, which connects the focal point, the top of the object and a point on the mirror, is reflected parallel to the optical axis. (It is just the reverse of ray B, described in Figure 5.)

Figure 7 shows the formation of a real image by a concave mirror when the object is located between the focal point and the center of curvature.

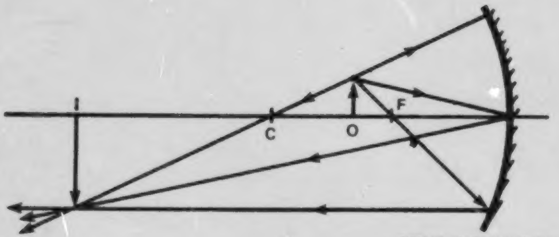


FIG. 7

Shown in Figure 8 is the formation of a virtual image by a convex mirror. As in a plane mirror, the image is the apparent origin of the reflected rays. This image is the only type produced by convex mirrors.

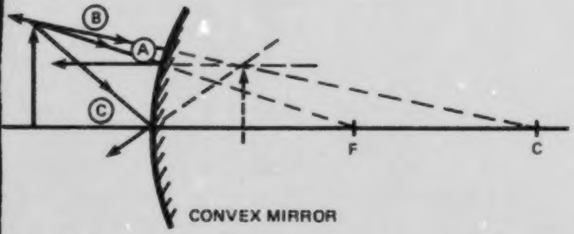


FIG. 8

IMAGE LOCATION AND SIZE

The location of an image formed by any mirror may be calculated by using Equation 1:

$$\frac{1}{f} = \frac{1}{p} + \frac{1}{q} \qquad (1)$$

where

f = focal length.

p = object distance.

q = image distance.

In this equation, p is the distance from the object to the mirror surface along the optical axis, and q is the distance from the image to the mirror surface.

Concave mirrors have a positive focal length and convex mirrors have a negative focal length. The image distance is positive for real images and negative for virtual images. Examples A and B show the use of Equation 1 for solving problems.

EXAMPLE A: IMAGE DISTANCE WITH A CONCAVE MIRROR

Given: Focal length of mirror: $f = 25$ cm; object distance: $p = 75$ cm.

Find: Image distance (is it real or virtual?).

Solution:

$$\frac{1}{f} = \frac{1}{p} + \frac{1}{q}$$

$$\frac{1}{25\text{ cm}} = \frac{1}{75\text{ cm}} + \frac{1}{q}$$

$$\frac{1}{q} = \frac{1}{25} - \frac{1}{75} = \frac{3}{75} - \frac{1}{75} = \frac{2}{75}$$

$$q = \frac{75}{2}$$

$q = 37.5$ cm (a real image is formed because q is positive.)

EXAMPLE B: IMAGE DISTANCE WITH A CONVEX MIRROR

Given: Focal length of mirror: $f = -10$ cm; object distance: $p = 25$ cm.

Find: Image distance (is it real or virtual?)

Solution.

$$\frac{1}{f} = \frac{1}{p} + \frac{1}{q}$$

$$\frac{1}{-10\text{ cm}} = \frac{1}{25\text{ cm}} + \frac{1}{q}$$

$$\frac{1}{q} = -\frac{1}{10\text{ cm}} - \frac{1}{25\text{ cm}}$$

$$= -0.1\text{ cm}^{-1} - 0.04\text{ cm}^{-1}$$

$$\frac{1}{q} = -0.14\text{ cm}^{-1}$$

$$\therefore q = \frac{1}{-0.14\text{ cm}^{-1}}$$

$q = -7.14$ cm (virtual image)

Size of the image may be calculated by using Equation 2.

$$h_i = h_o \left(\frac{q}{p}\right) \qquad (2)$$

where

h_i = image height.
h_o = object height.
q = image distance.
p = object distance.

Ratio q/p is the magnification of the image, or h_i/h_o.

The use of Equation 2 is shown in Example C to solve a problem.

EXAMPLE C: IMAGE HEIGHT	
Given:	The object in Example A is 5 cm high; object distance: $p = 75$ cm; image distance: $q = 37.5$ cm.
Find:	Image height.
Solution:	$h_i = h_o \left(\frac{q}{p}\right)$ $= (5\text{ cm}) \times \left(\frac{37.5\text{ cm}}{75\text{ cm}}\right)$ $= 5\text{ cm} \times 0.5$ $h_i = 2.5$ cm (Magnification is 1/2.)

APPLICATION OF MIRRORS

Mirrors are employed in a wide variety of optical instruments and devices including telescopes, interferometers and lasers. Mirrors also are an essential part of many more common devices. Table 1 lists several mirror applications along with the ray diagram that applies to each application. Information on image distance and type and size of image also is included.

TABLE 1. APPLICATIONS OF IMAGING WITH MIRRORS

Mirror Type	Object Distance	Image Distance	Image Type/Size	Ray Diagram	Application
Plane	Any	$q = -p$ (As far behind as object is in front)	Virtual/ same size as object	Figure 4	Ordinary mirror
Convex	Any	$q < f$	Virtual/ reduced	Figure 8	Wide-angle rear vision mirror Antitheft mirror in store
Concave	$p = \infty$	$q = f$	Real/ reduced to point	Figure 2	Solar furnace (parallel incident rays)
Concave	$p = f$	$q = \infty$	None	Figure 2	Spot light (parallel reflected rays)
Concave	$p < f$	$q < 0$	Virtual/ enlarged	Figure 6	Magnifying mirror
Concave	$f \leq p \leq 2f$	$q \geq 2f$	Real/ enlarged	Figure 7	Image projection system of large screen television
Concave	$p > 2f$	$2f > q \geq f$	Real/ reduced	Figure 5	Reflecting telescope

SUMMARY

Mirrors change the direction of light rays by reflection. Plane mirrors reflect parallel incident rays as parallel rays. Concave mirrors reflect parallel rays so that they converge at a point (called the focal point of the mirror). Convex mirrors reflect parallel rays as though they originated at a focal point behind the mirrors' surfaces. The location, type and size of the image produced by a mirror may be determined by drawing two or more rays reflected from the mirror surface according to the law of reflection.

IMAGE FORMATION WITH CONCAVE MIRROR

MATERIALS

Concave mirror 16-inch diameter, 22-inch focal length (Edmund #80,097 or equal).
Piece of wood approximately 1" x 1/4" x 12".
Meter Stick.
Shoe Box.
Tape.
Coin.

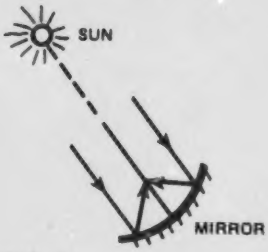


FIG. 9

PROCEDURE

1. Place concave mirror in direct sunlight with optical axis of mirror directed toward sun (Figure 9). CAUTION: Do not place hand or clothing near focal point of mirror.

2. Use piece of wood to locate focal point of mirror. Parallel rays of sun will be focused to smallest spot at focal point. In bright sunlight, wood will burst into flames, demonstrating collection of solar energy. Measure focal length of mirror with meter stick; avoid positioning hand or meter stick at focal point; record focal length in Data Table.

3. Suspend coin inside open end of shoe box; place open end of box in front of mirror, as shown in Figure 10. Top edge of coin is on optical axis; object distance equals radius of curvature (twice focal length).

4. Observe real image produced. Bottom edge of coin is on optical axis. With proper placement of coin, image will appear to stand on edge of box.

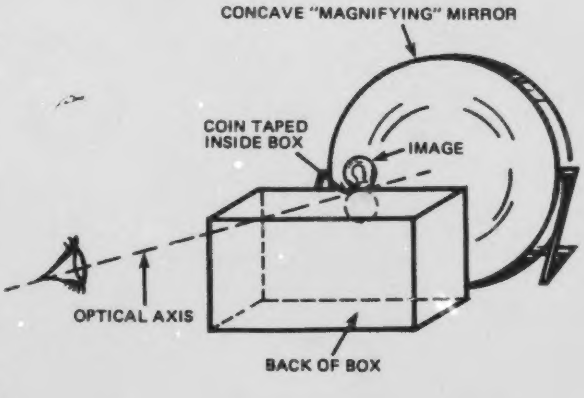


FIG. 10

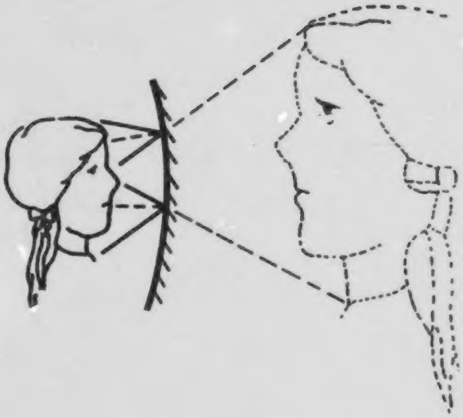


FIG. 11

5. Complete ray diagram of real image formation in Data Table.

6. Hold concave mirror 10 to 12 inches in front of face (Figure 11); observe virtual image formed in mirror.

7. Complete ray diagram of virtual image formation in Data Table.

8. How are real and virtual images utilized in practical applications?

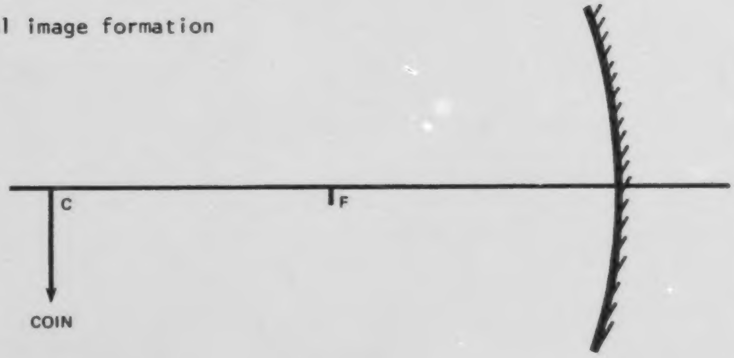
DATA TABLE

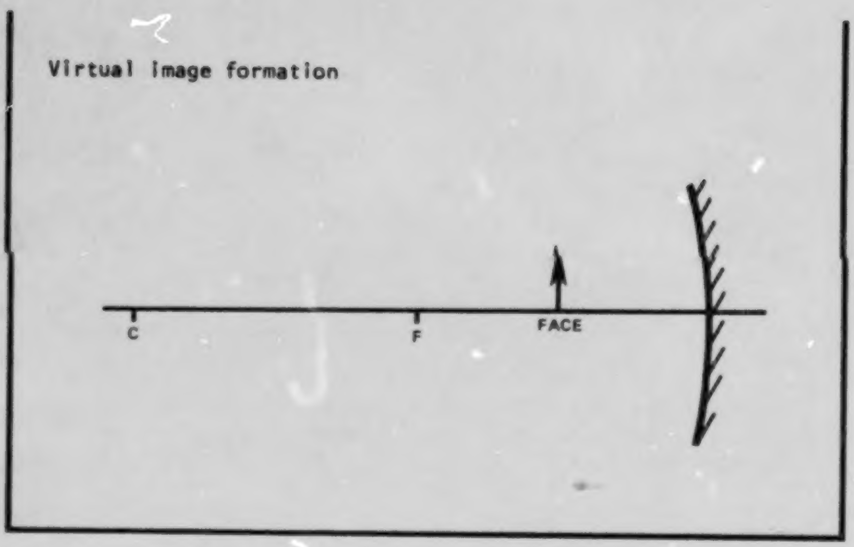
Focal length of mirror.

Specified: f = _____

Measured: f = _____

Real image formation





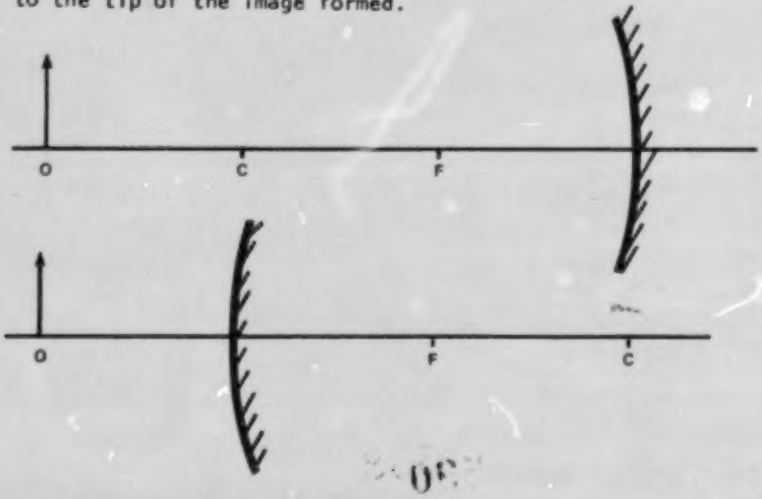
Virtual image formation

SELF CHECK

1. Describe difference between a real image and a virtual image.

2. Draw diagrams to illustrate the focal points of convex and concave mirrors; show parallel light rays reflected from these mirrors.

3. On the following diagrams, draw three rays from the tip of the object to the tip of the image formed.



4. A concave mirror has a focal length of 10 cm. An object is located 15 cm from this mirror. Where is the image?

5. The object in problem 4 is 2 cm high. What is the image height?

6. An object is located 10 cm from a mirror. A virtual reduced, image is formed 2 cm behind the mirror surface. What is the focal length of the mirror? Is it concave or convex?

7. The image in problem 6 is 0.6 cm high. What is the object height?

8. Parallel light is incident on a concave mirror that has a radius of curvature of 30 cm. Where will the light be brought to a focus?

13L9

PRISMS — REFRACTION AND DISPERSION OF LIGHT

In a vacuum, light travels at 3×10^8 m/s. When light radiation travels through any transmitting material, its speed always is lower than that in a vacuum. This reduction of speed inside materials is the basis of refraction (bending) of light and of the dispersion (spreading of colors) by prisms. This module presents refraction of light at plane surfaces and a discussion of variations in angle of refraction for different colors of light. In the laboratory, the student will use a prism spectrometer to measure light velocity and index of refraction of a prism at several wavelengths.

OBJECTIVES

Upon completion of this module, the student should be able to:

- State the equation for the law of refraction, and illustrate with a diagram.
- Given any two of the following quantities, calculate the third:

    Speed of light in a vacuum.
    Speed of light in a material.
    Index of refraction of the material.

- Given any three of the following quantities, calculate the fourth:

    Angle of incidence.
    Angle of refraction.
    Index of refraction of first material.
    Index of refraction of second material.

- Explain total internal reflection in a short paragraph, including a prism as an example.
- Given the appropriate equipment, measure light velocity and index of refraction of a prism at several wavelengths.

DISCUSSION

INDEX OF REFRACTION

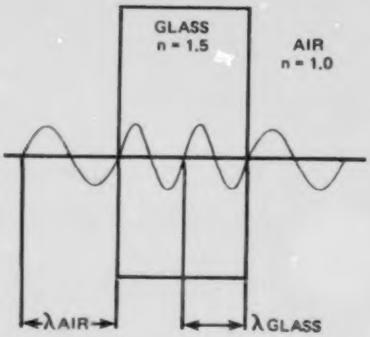


FIG. 1

Figure 1 shows a light wave traveling through a piece of glass. As the light wave enters the glass, it slows down, and its wavelength becomes shorter. When the light leaves the glass, both its speed and wavelength increase to their original values. (Frequency always remains constant.) The index of refraction of a material is the ratio of the speed of light in a vacuum to its speed in that material, as defined by Equation 1.

$$n = \frac{c}{v} \qquad (1)$$

where

n = index of refraction.
c = speed of light in a vacuum.
v = speed of light in material.

The index of refraction of a vacuum is 1, but is greater than one for other materials.

Example A shows the use of Equation 1 in solving a problem.

EXAMPLE A: SPEED OF LIGHT IN A DIAMOND

Given: Speed of light in vacuum: $c = 3 \times 10^8$ m/s.
Index of refraction of diamond: $n_d = 2.4$.

Find: Speed of light in a diamond.

Solution:

$$v = \frac{c}{n}$$

$$= \frac{3 \times 10^8 \text{ m/s}}{2.4}$$

$$v = 1.25 \times 10^8 \text{ m/s}$$

The index of refraction of air may be assumed to be 1.0 for most applications.

REFRACTION OF LIGHT

Refraction is the change in direction of travel of light as it crosses the boundary between two materials.

Figure 2 shows the refraction of light at the surface of a piece of glass. The "normal line" (Figure 2b) is a line perpendicular to the optical surface, and all angles are measured from the normal. Angle A is called the "angle of incidence," and Angle C is called the "angle of refraction." The dashed lines in Figure 2a represent wavefronts (surfaces of constant phase). As the light enters the glass, it slows down, and its wavefronts become closer together and change direction as shown. The direction of travel of the light is "bent" toward the normal according to the law of refraction given in Equation 2. Note that wavelength in the glass is shorter than in air.

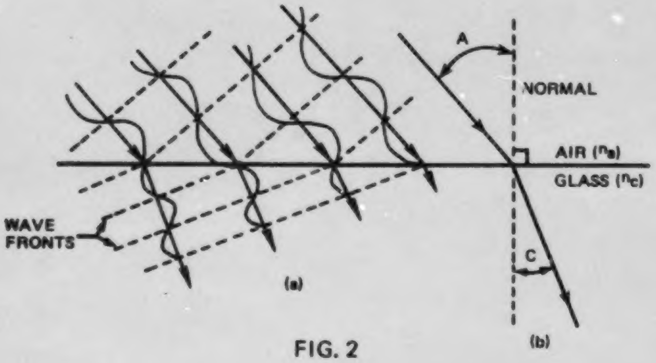


FIG. 2

$$n_a \sin A = n_c \sin C \qquad (2)$$

where

n_a = index of refraction of first material.
$\sin A$ = sine of angle of incidence.
n_c = index of refraction of second material.
$\sin C$ = sine of angle of refraction.

Examples B and C show the use of this equation in solving problems.

EXAMPLE B: ANGLE OF REFRACTION

Given: In Figure 2b angle of incidence: $A = 42°$.
Index of refraction of air: $n_a = 1.0$.
Index of refraction of glass: $n_c = 1.7$.

Find: Angle of refraction C.

Solution: $n_a \sin A = n_c \sin C$

$$(1.0)\ \sin\ (42°) = (1.7)\ \sin\ C$$

$$(1.0) \times (0.67) = (1.7)\ \sin\ C$$

$$\sin\ C = \frac{(1.0)\ (0.67)}{(1.7)}$$

$$\sin\ C = 0.394$$

$$C = \sin^{-1}\ (0.394)$$

$$C = 23.2°$$

EXAMPLE C: REFRACTION BY A PRISM

Given: Light refracted by a 60° prism. The angle of incidence at entrance is 70°, and the angle of refraction is 37°. The angle of incidence at exit is 23°.

Find:

(a) Index of refraction of prism (n_p).

(b) Angle of refraction at exit (F).

Solution:

(a)

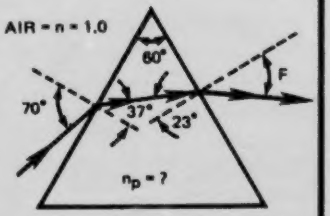
$$n_a\ \sin\ A = n_p\ \sin\ C$$

$$n_p = \frac{n_a\ \sin\ A}{\sin\ C}$$

$$n_p = \frac{(1.0)\ \sin\ (70°)}{\sin\ (37°)}$$

$$= \frac{0.940}{0.602}$$

$$n_p = 1.56$$



(b)

$$\sin\ C = \frac{n_a\ \sin\ A}{n_c}$$

$$\sin\ F = \frac{(1.56)\ \sin\ (23°)}{1.0} \text{ (See note below.)}$$

$$= \frac{(1.56)(0.391)}{1.0}$$

$$\sin\ F = 0.610$$

$$F = \sin^{-1}\ (0.61)$$

$$F = 37.6°$$

Notice that in part b the light is incident upon air from inside the glass prism. Thus: $n_a = 1.56$, $A = 23°$, $n_c = 1.0$, $C = F$.

The refraction of light by prisms has many applications in lighting. Figure 3 shows an ordinary plastic diffuser used in both fluorescent and incandescent light fixtures to produce more even illumination. The tiny prisms change direction of

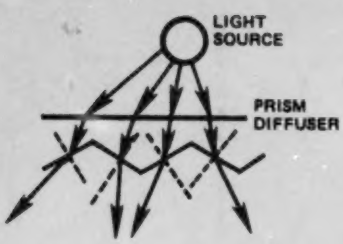


FIG. 3

the light rays and increase apparent size of the light source.

TOTAL INTERNAL REFLECTION

Example D shows what happens when light strikes an air-glass boundary from inside the glass with a high angle of incidence.

EXAMPLE D: TOTAL INTERNAL REFLECTION

Given: Light strikes an air-glass boundary from inside the glass (n = 1.5) with an angle of incidence of 50°.

Find: Angle of refraction C.

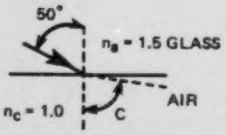
Solution: $n_a \sin A = n_c \sin C$

$$\sin C = \frac{n_a \sin A}{n_c}$$

$$= \frac{(1.5)\sin(50°)}{(1.0)}$$

$$= (1.5)(0.766)$$

$$\sin C = 1.149$$



Something is wrong, because the sine of an angle can never be greater than 1. This problem cannot be solved by law of refraction. (Refer to next paragraph and to Figure 4.)

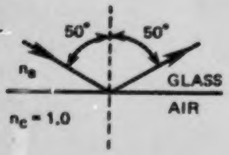


FIG. 4

Figure 4 shows what actually happens in Example D. The light cannot be refracted out of the glass, but experiences total internal reflection at the surface. The greatest angle of incidence for which light can be refracted through a boundary is called the "critical angle," and is given by Equation 3.

$$A_c = \sin^{-1}\left(\frac{n_a}{n_c}\right) \tag{3}$$

where

A_c = critical angle.

$\sin^{-1}$ = arc sine.

If the angle of incidence exceeds the critical angle, all the light is reflected according to the law of reflection discussed in Application Module 13L7, "Specular and Diffuse Reflection of Light." Figure 5 shows how this principle is used in reflective prisms. The plastic reflectors along the sides of roads and on automobiles are composed of many small prisms that reflect light back in the direction from which it came. Road signs are seen easily at night because they are covered with tiny glass beads that serve the same function. Figure 6 shows light reflected by such a glass sphere. In this case, the angle of incidence is less than the critical angle, and some of the light is refracted out of the rear of the sphere.

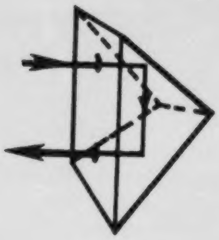


FIG. 5

DISPERSION

Figure 7 shows how a prism separates white light into the visible spectrum by a process called "dispersion." This phenomenon occurs because the index of refraction of materials varies with the wavelength of light, which indicates that different colors (wavelengths) of light have different speeds through a material.

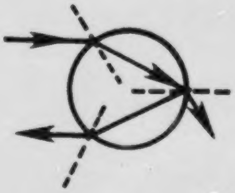


FIG. 6

Figure 8 shows dispersion curves for several optical materials. In each case, the index of refraction is highest for shorter wavelengths. Thus, shorter wavelengths (blue light) are refracted to a greater extent than longer wavelengths (red light). The ability of a prism to disperse white light into its component wavelengths is the basis of the prism spectrometer used in the laboratory portion of this module. The same principle is employed in infrared spectrophotometers used for analysis in scientific laboratories.

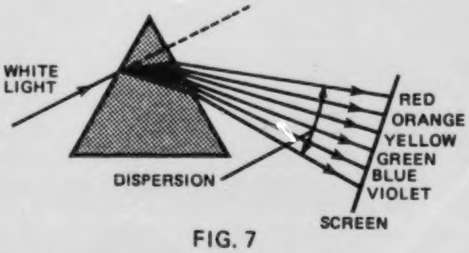


FIG. 7

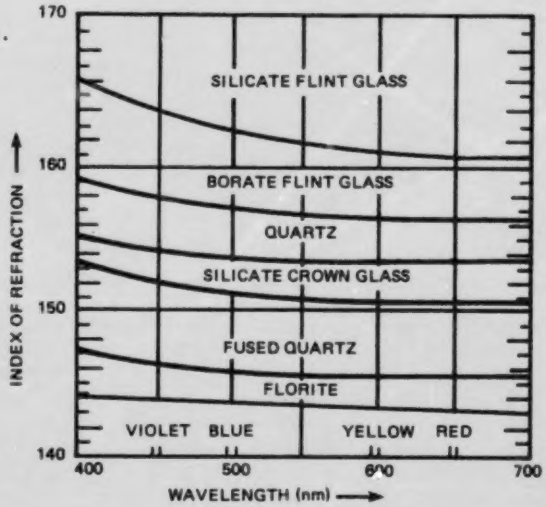


FIG. 8

SUMMARY

Reduction of light speed when passing through a prism is responsible for light refraction (bending) and for light dispersion (spreading of colors). Prisms are used in many reflective materials. Prism and infrared spectrometers disperse white light into its component wavelengths, making visible the various colors.

LIGHT VELOCITY AND INDEX OF REFRACTION OF A PRISM

MATERIALS

Prism spectrometer (Sargent-Welch #S-75903-80 or equivalent consisting of:

- Collimator with slit and lens that allows a small amount of light to strike prism in parallel rays.
- Equilateral prism.
- Telescope with objective lens, eyepiece and cross hairs.
- Vernier scale to allow measurement of angles to the nearest tenth of a degree.

Helium spectrum tube.
Spectrum-tube power supply.

PROCEDURE

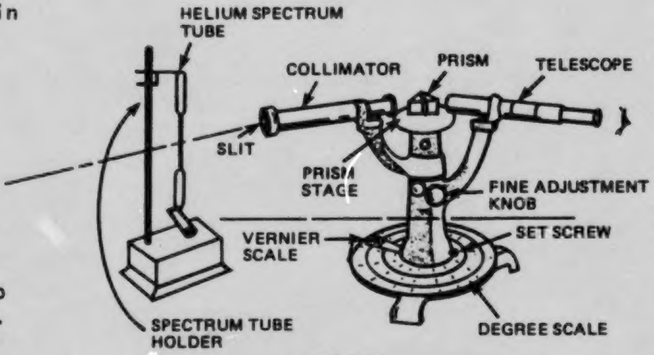

HELIUM SPECTRUM TUBE
COLLIMATOR
PRISM
TELESCOPE
SLIT
PRISM STAGE
FINE ADJUSTMENT KNOB
VERNIER SCALE
SET SCREW
SPECTRUM TUBE HOLDER
DEGREE SCALE

FIG. 9

1. Set up equipment as shown in Figure 9, but do not place prism on prism stage.
2. Turn on spectrum tube power supply; place spectrometer slit near capillary portion of tube. BEWARE OF HIGH VOLTAGE.
3. Align telescope with collimator; focus telescope to obtain sharp image of slit. Set cross hair on image of slit.
4. Loosen set screw; move degree scale around until zero on degree scale lines up with zero on vernier scale. Be sure cross hair is on image of slit; tighten set screw.
5. Place prism on stage as in Figure 9. Move telescope to angle of approximately 40°. Helium tube and collimator remain fixed.

6. View spectral lines. Refocus telescope for sharpest image.

7. Rotate prism stage while observing spectrum. Set prism to obtain minimum angle of deviation of light. NOTE: The angle of deviation is the angle measured between the straight line found by the collimator and telescope (step 3) and the telescope position after the prism has been placed on the stage (step 5). This angle can be made a minimum for any line in the helium spectrum viewed with the telescope when the prism is rotated appropriately. Consult instructor if difficulty arises with this setting.

8. Set cross hair on first red line (706.8 nm). Read angle from vernier; record in Data Table.

9. Without moving prism, set cross hair on each bright helium line listed in Data Table; record angle of each.

10. Calculate index of refraction for prism at each wavelength using the following equation (true only at angle of minimum deviation):

$$n = \frac{\sin\left(\frac{A + B}{2}\right)}{\sin (A/2)}$$

where

A = apex angle of prism (60°).
B = angle of deviation of light.

Record index of refraction to five significant figures.

11. Choose appropriate values for vertical axis of graph (Figure 10) to show variations in index of refraction. Plot data points; connect with a smooth curve.

DATA TABLE			
Helium Spectrum		Angle of Deviation °	Index of Refraction
Color	Wavelength (nm)		
Red	706.8		
Red	667.8		
Yellow	587.6		
Green	504.8		
Blue-green	492.2		
Blue	471.3		
Deep blue	447.1		
Violet	438.8		
Violet	402.6		
Violet	388.6		

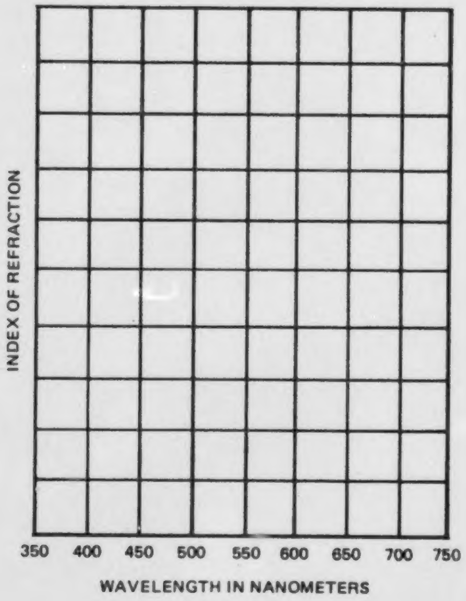


FIG. 10

SELF CHECK

1. Write the equation for the law of refraction.

2. The index of refraction of fused quartz for light with a wavelength of 700 nm is 1.45. What is the speed of light in quartz?

3. Light passes from air into water (n = 1.33) with an angle of refraction of 32°. What is the angle of incidence?

4. Light passes from water (n = 1.33) into flint glass (n = 1.64) with an angle of incidence of 55°. What is the angle of refraction.

5. Draw a sketch showing total internal reflection in a prism.

6. Is the index of refraction of a material greater for red light or blue light?

7. Find the index of refraction of a material in which light has a speed of 2×10^8 m/s.

8. Light is incident upon a water (n = 1.33) – glass (n = 1.5) boundary. What is the greatest angle of incidence for which light can be refracted through this boundary?

13L10

LENSES — FOCUSING AND SPREADING OF LIGHT

Lenses are optical devices that change the directionality of light beams by refracting different parts of the beams by differing amounts. This module discusses properties of lenses, formations of real and virtual images by lenses, and a variety of lens applications, including corrective eyewear. In the laboratory, the student will construct a small refracting telescope.

OBJECTIVES

Upon completion of this module, the student should be able to:

- Draw an appropriate diagram that illustrates each of the definitions of image forming systems given below:

  The focal point of a positive lens.
  The focal point of a negative lens.
  The real image formed by a positive lens.
  The virtual image formed by a positive lens.
  The virtual image formed by a negative lens.

- Give one practical application for each of the above optical arrangements.
- Given any two of the following quantities, calculate the third:

  Focal length of lens.
  Object distance.
  Image distance.

- Given one of the following quantities, calculate the other:

  Focal length of lens.
  Power of lens in diopters.

- Given appropriate equipment, form and describe real and virtual images developed by positive lenses; construct a refracting telescope.

FOCAL POINTS OF LENSES

In Figure 1, parallel rays of light strike a converging (positive) lens. Each ray is refracted at both surfaces of the lens according to the law of refraction (discussed in Application Module 13L9, "Prisms - Refraction and Dispersion of Light"). The rays then converge to point F, the focal point of the positive lens. Distance f, from the center of the lens to the focal point, the focal length of the lens, is a positive quantity for a positive lens. Practical applications of this ray diagram include 1) focusing of the sun's parallel rays or 2) collimating the beam from a laser.

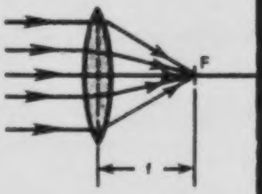


FIG. 1

Figure 2 shows the effect of a diverging (negative) lens on parallel light rays. The lens acts to diverge parallel rays as though they had originated (dashed lines) from a single point F (the focal point of the negative lens). Focal length f of the negative lens is a negative quantity. This ray diagram, which represents the spreading of light rays to cover a larger area, often is used in photographic flash systems.

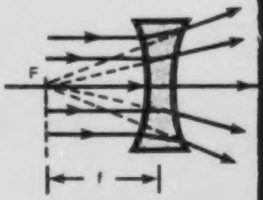


FIG. 2

The focal length of a lens may be measured in any convenient units of length; often, this refractive power of a lens is expressed in "diopters," as expressed in Equation 1.

$$P\ (D) = \frac{1}{f\,(m)}. \qquad (1)$$

where

P = refractive power in diopters.
f = focal length in meters.

Example A shows how to use Equation 1 in solving a typical problem.

EXAMPLE A: POWER OF A LENS

Given:
(a) Focal length of a converging lens is 25 cm.
(b) Power of a diverging lens is -2.5.

Find:
(a) Power of converging lens.
(b) Focal length of negative lens.

Solution:

(a)
$$P = \frac{1}{f} = \frac{1}{0.25\ \text{m}}$$
$$P = 4.0\ \text{D}$$

(b)
$$f = \frac{1}{P} = \frac{1}{-2.5\ \text{D}}$$
$$f = -0.4\ \text{m} = -40\ \text{cm}$$

Figure 3 shows formation of a real image by a positive lens. Object 0 is at a distance from the lens greater than twice the focal length (2f) of the lens. The following rays are used to locate the image:

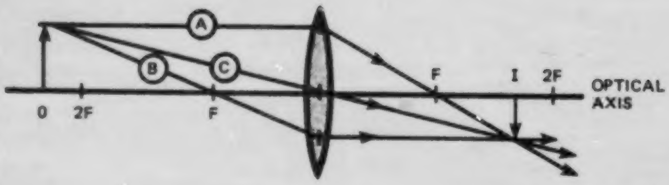


FIG. 3

Ray A originates at the top of the object and travels parallel to the optical axis. It is refracted to pass through the focal point on the other side of the lens.

Ray B travels through the focal point on the object side of the lens and is refracted to leave the lens in a direction parallel to the optical axis.

Ray C passes through the center of the lens and is undeviated.

If the object is located at a great distance from the lens, the image is near the focal point and greatly reduced in size. This optical arrangement applies both to a camera and to the human eye.

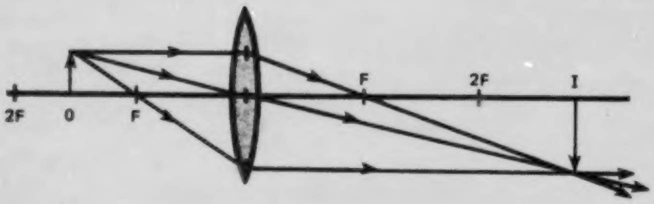


FIG. 4

Figure 4 shows the image formed by the same lens when the object distance is in the range between the focal length of the lens and twice the focal length. In this instance, the image is enlarged and the image distance is greater than 2f. This ray diagram is representative of the optical arrangements used in movie and slide projectors, where 0 is the film and I is the screen.

Figure 5 shows the virtual image formed when an object is placed inside the focal point of a positive lens. This virtual image always appears magnified and erect. (Real images formed by lenses always appear inverted.) Virtual, magnified images formed by positive lenses include those displayed by the familiar magnifying glasses, as well as the plastic lenses used to enlarge the display of many pocket calculators.

The virtual image formed by a negative lens is shown in Figure 6. In this diagram, ray A enters the lens parallel to the optical axis (as in the case of the positive lens), and is refracted as though it came from the focal point. Ray B travels toward the focal point on the opposite side of the lens, and is refracted parallel to the optical axis. Ray C passes through the center of the lens and remains undeviated. The virtual image formed is always erect and smaller than the object. Peepholes in doors often are equipped with negative lenses, which provide wide-angle views.

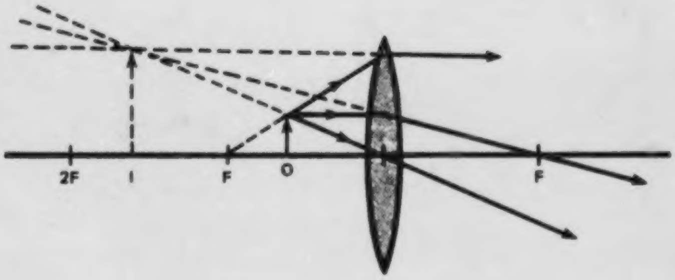


FIG. 5

IMAGE LOCATION AND SIZE

Equations for location and size of images formed by lenses are the same as those for mirrors given in Module 13L8, "Image Formation with Mirrors":

$$\frac{1}{f} = \frac{1}{p} + \frac{1}{q} \qquad (2)$$

where
f = focal length.
p = image distance.
q = object distance.

$$h_i = h_o\left(\frac{q}{p}\right) \qquad (3)$$

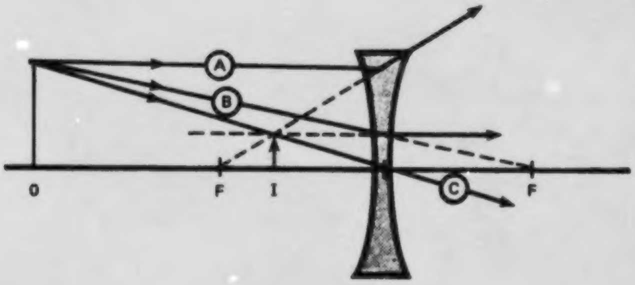


FIG. 6

where

h_i = image height.
h_o = object height.
q = image distance.
p = object distance.

The ratio q/p is the magnification of the lens. Equation 3 shows that this is the same as the ratio of image size to object size, or h_i/h_o.

Examples B and C illustrate the use of Equations 2 and 3 in solving problems.

EXAMPLE B: IMAGE FORMATION WITH A CONVERGING LENS

Given: Focal length of lens: $f = 10$ cm; object distance: $p = 25$ cm; object height: 3 cm.

Find: Image distance and height.

Solution: Image distance =

$$\frac{1}{f} = \frac{1}{p} + \frac{1}{q}$$

$$\frac{1}{10\text{ cm}} = \frac{1}{25\text{ cm}} + \frac{1}{q}$$

$$0.1\text{ cm}^{-1} = 0.04\text{ cm}^{-1} + \frac{1}{q}$$

$$\frac{1}{q} = 0.06\text{ cm}^{-1}$$

$$q = \frac{1}{0.06\text{ cm}^{-1}}$$

$$q = 16.67\text{ cm}$$

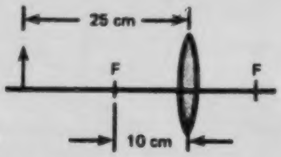


Image height =

$$h_i = h_o \times \left(\frac{q}{p}\right)$$

$$= 3.0\text{ cm} \times \left(\frac{16.67\text{ cm}}{25\text{ cm}}\right)$$

$$h_i = 2.0\text{ cm}$$

Magnification of lens = $h_i/h_o = 2/3 = q/p = 16.67/25$

EXAMPLE C: IMAGE FORMATION WITH A DIVERGING LENS

Given: Object distance: $p = 10$ cm; object height: $h_o = 3.5$ cm; virtual height: $h_i = 0.625$ cm.

Find: Image distance and focal length of lens.

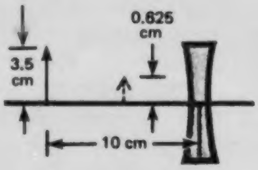
Solution: Image distance=

$$h_i = h_o \left(\frac{q}{p}\right)$$

$$q = p \left(\frac{h_i}{h_o}\right)$$

$$q = (10\text{ cm}) \times \left(\frac{0.625\text{ cm}}{3.5\text{ cm}}\right)$$

$$q = 1.79\text{ cm}$$



Focal length =

$$\frac{1}{f} = \frac{1}{p} + \frac{1}{q}$$

$$= \frac{1}{10\text{ cm}} + \frac{1}{(-1.79\text{ cm})}$$ image distance of virtual image is negative.

$$= 0.1\text{ cm}^{-1} - 0.559\text{ cm}^{-1}$$

$$= -0.459\text{ cm}^{-1}$$

$$\therefore f = \frac{1}{-0.459\text{ cm}^{-1}}$$

$$f = -2.18\text{ cm}$$

LENS COMBINATIONS

Many optical systems consist of several lenses in combination. Figure 7 shows a sixteen-element lens system, such as a zoom lens for a movie camera. Figure 8a shows proper focusing of parallel light rays by an eye; Figure 8b shows focusing of parallel light by a nearsighted (myopic) eye; Figure 8c shows that of a farsighted (hyperopic) eye. In Figure 9, corrections for myopia (9a) and hyperopia (9b) are achieved by adding the appropriate correcting lens to an optical system.

Figure 10 shows the locations of images in a microscope. The objective lens forms a real image of object A at point B. This real image serves as the object of the eyepiece lens, which forms a virtual image at C. The lens of the eye, which uses the virtual image at C as an object, focuses a real image on the retina at D.

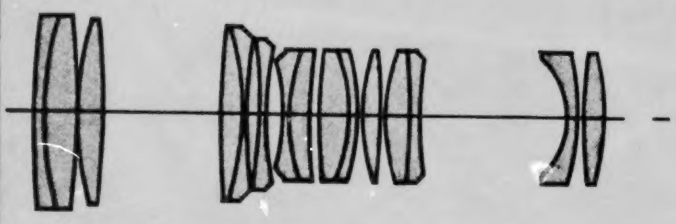


FIG. 7

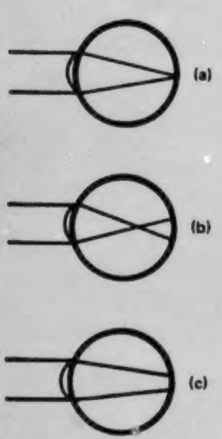


FIG. 8

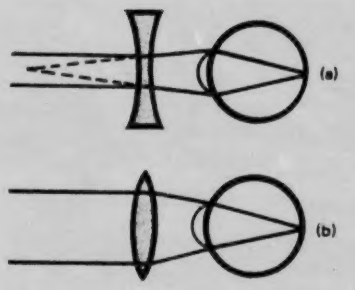


FIG. 9

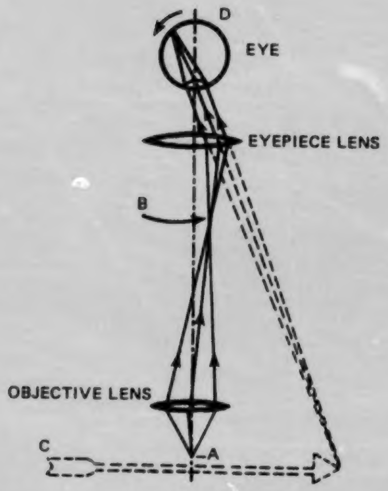

D
EYE
EYEPIECE LENS
B
OBJECTIVE LENS
C
A

FIG. 10

SUMMARY

A lens changes the directionality of a light beam by refracting different parts of the beam through differing angles. A converging (positive) lens is one that causes parallel incident rays to converge to a point (called the focal point of the lens). A diverging (negative) lens is one that causes parallel rays to diverge as though they originated from the focal point. The size, type and location of an image formed by a lens may be determined by drawing two or more rays through the lens.

IMAGE FORMATION WITH LENSES

MATERIALS

Refractor telescope kit (Edmund #71,473 or equivalent) consisting of:

- Objective lens.
- Eyepiece lens.
- Glare stops.
- Cardboard tubes.
- End caps.
- Instructions for assembly.

Incandescent light bulb.
White screen.

PROCEDURE

1. Set incandescent bulb and screen 1 meter apart on level surface (Figure 11).
2. Hold objective lens (larger-diameter lens) near screen and perpendicular to light rays from bulb; move lens toward bulb until sharp image is formed on screen.
3. Measure image (q) and object (p) distances; record in Data Table.
4. Complete Data Table under "Objective Lens" to determine focal length of objective lens.
5. Repeat above procedure for eyepiece lens.

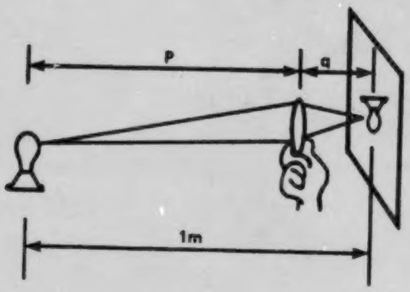


FIG. 11

6. Assemble telescope according to directions provided with telescope kit.
7. View virtual images of distant objects formed by telescope.
8. Complete ray diagram of telescope in Data Table below.

DATA TABLE

Objective lens

Object distance: $p = \_\_\_\_\_\_$ cm.

Image distance: $q = \_\_\_\_\_\_$ cm.

Focal length: $\frac{1}{f} = \frac{1}{p} + \frac{1}{q}$

$$= \frac{1}{\_\_\_\_\ \text{cm}} + \frac{1}{\_\_\_\_\ \text{cm}}$$

$$= \_\_\_\_\_\ \text{cm}^{-1} + \_\_\_\_\_\ \text{cm}^{-1}$$

$$= \_\_\_\_\_\ \text{cm}^{-1}$$

$$f = \frac{1}{\_\_\_\_\ \text{cm}^{-1}}$$

and $f = \_\_\_\_\_$ cm

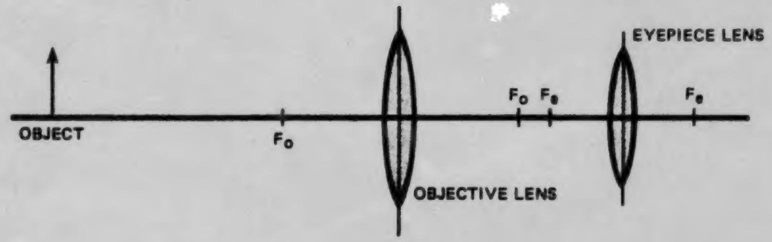
Eyepiece lens

Object distance: $p = \_\_\_\_\_$ cm

Image distance: $q = \_\_\_\_\_$ cm

Focal length: $f = \_\_\_\_\_$ cm

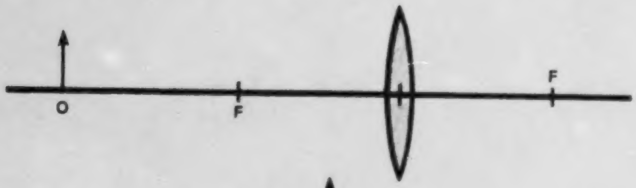
Ray diagram of a telescope:


OBJECT
F_o
OBJECTIVE LENS
F_o F_e
EYEPIECE LENS
F_e

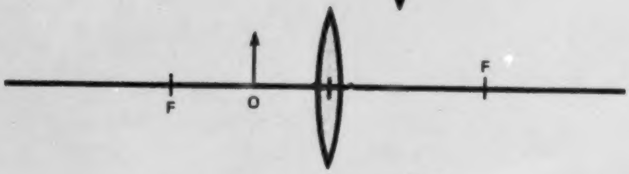
SELF CHECK

1. Complete the following ray diagrams to show the image formed.

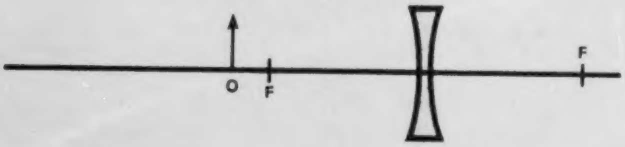
a.



b.



c.



2. Sketch the lens, object and image in each of the following optical systems:

   a. Magnifying glass.
   b. Movie projector.
   c. Camera.

3. The focal length of a peephole lens in a door is -2.0 cm. A person 6 feet tall stands 3 m from the lens; find the image distance and height.

4. A slide projector lens has an object distance of 8.5 cm and an image distance of 2.3 m. What is its focal length?

5. What is the focal length of a lens with a power of 10 diopters?

6. What is the power in diopters of a lens with a 2.5 cm focal length?

13L11

INTERFERENCE — TESTING FLAT SURFACES

The interference of light waves is a description of the way in which light radiation interacts with itself. If two light waves are in phase at a point on a screen, a bright spot is produced. If the two waves are completely out of phase, a dark area results. The alternating light and dark fringes of such an interference pattern can provide useful information. This module discusses the interference of light waves and one method of using this phenomenon for highly-accurate measurement of the flatness of surfaces. In the laboratory, the student will produce air-wedge interference patterns.

OBJECTIVES

Upon completion of this module, the student should be able to:

- Draw and itentify diagrams that illustrate constructive and destructive interference of waves.
- Draw and identify a diagram that depicts the light paths for the constructive and destructive interference of light waves reflected from an air wedge.
- Given the wavelength of light and the spacing of the fringes produced by an air wedge, calculate the wedge angle.
- Illustrate the interference pattern formed when each of the following is placed on a flat surface:

  Another flat surface at a small wedge angle.

  Another flat surface at a large wedge angle.

  A spherical surface.

- Given the appropriate equipment, produce and sketch air-wedge interference patterns.

DISCUSSION

Figure 1a depicts two light waves with their electric fields in phase, arriving at a point on the screen. The two electric fields add together to form a resultant wave (dashed line), of which the amplitude is the sum of the amplitudes of the component waves. The irradiance in Figure 1a is four times the irradiance produced by one wave alone, because irradiance increases with the square of the amplitude of the electric field. This process is called "constructive interference," which results in a bright spot of light.

Figure 1b illustrates destructive interference. In this figure, two light waves arrive exactly out of phase, and their electric fields cancel to produce a dark spot. Interference actually is a redistribution of the optical energy present. That energy missing in the destructive interference areas (dark) appears in the constructive interference areas (bright).

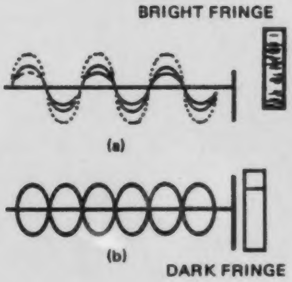

BRIGHT FRINGE
(a)
(b)
DARK FRINGE

FIG. 1

INTERFERENCE PATTERNS

Figure 2 shows how a pattern of light and dark areas (called an "interference pattern") can be produced. Two flat pieces of glass are separated by a thin air wedge and illuminated by a monochromatic (single-color) light source. Under the proper conditions, a series of light and dark bands, called "fringes," will be visible to an observer looking down onto the glass plates.

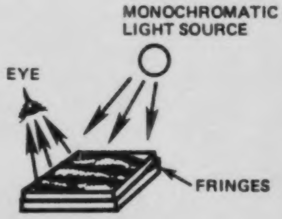

MONOCHROMATIC
LIGHT SOURCE
EYE
FRINGES

FIG. 2

Figure 3 indicates how the interference pattern is formed. This figure shows the air wedge between the two pieces of glass and the reflections from the two glass surfaces. Each surface reflects only about four percent of the light striking it; so the patterns produced are dim and must be viewed against a dark background. In Figure 3a, the reflection from the upper and lower surfaces on the air wedge reach the eye exactly out of phase, to create a dark fringe. In Figure 3b, the two waves are reflected in phase to produce a light fringe. These two conditions alternate across the glass, to produce a series of light and dark fringes.

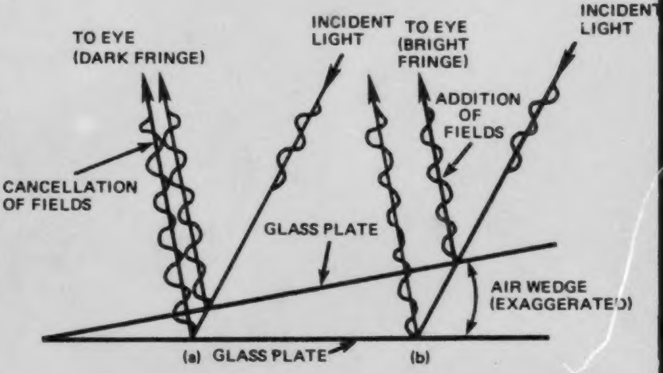

TO EYE
(DARK FRINGE)
INCIDENT
LIGHT
TO EYE
(BRIGHT
FRINGE)
INCIDEN
LIGHT
ADDITION
OF
FIELDS
CANCELLATION
OF FIELDS
GLASS PLATE
AIR WEDGE
(EXAGGERATED)
(a)
GLASS PLATE
(b)

FIG. 3

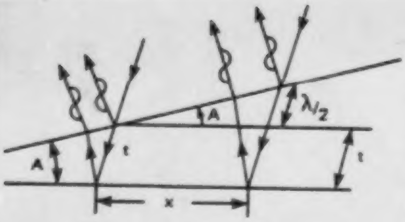


FIG. 4

Figure 4 represents two adjacent, bright fringes produced by an air wedge. The thickness of the wedge in Figure 4a is t; therefore, the wave reflected from the lower surface traves 2t farther than the wave reflected from the upper surface. At Figure 4b, the thickness of the wedge has increased by one-half the wavelength of the light used ($\lambda/2$), which means that the path length of the lower wave has increased by an amount equal to the wavelength of the light. Thus, the two reflected waves once again are in phase for constructive interference. Halfway between these two points, the two waves are out of phase, producing destructive interference. The wedge angle of such an air wedge can be calculated from Equation 1.

$$A = \frac{\lambda}{2x} \tag{1}$$

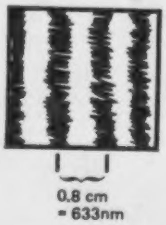
where

A = wedge angle in radians.
λ = wavelength of light.
x = fringe spacing.

The fringe spacing x is the linear distance between the center of one bright fringe and the center of the next bright fringe, or from dark fringe to dark fringe. Example A shows how this equation is used in solving a problem.

EXAMPLE A: WEDGE ANGLE CALCULATION

Given: An interference pattern exhibits a fringe spacing of 0.8 cm. The wavelength of light is 633 nm.



Find: Angle of air wedge producing fringe pattern.

Solution:

$$A = \frac{\lambda}{2x}$$

$$= \frac{6.33 \times 10^{-7}\ \text{m}}{(2)(8 \times 10^{-3}\ \text{m})}$$

$$A = 3.96 \times 10^{-5}\ \text{rad.}$$ (A small angle!)

FLATNESS OF OPTICAL SURFACES

The technique described in this module commonly is used to determine the flatness of optical surfaces by comparing them with a surface of known flatness, called a "master," or "standard." If the master is truly flat, the interference pattern produced by placing the unknown surface in contact with the master surface is a contour map of the air space between the two and, therefore, is a reliable indication of the shape of the unknown surface. This procedure is demonstrated in Figures 5 through 9. In each case, the lower reflecting surface is the master flat.

If both surfaces are flat, the interference pattern will be a series of straight, regularly-spaced fringes. Broad fringes (Figure 4) indicate a small angle between the surfaces, and fine fringes (Figure 5) indicate a larger angle; but both of these patterns signify a flat surface.

Figure 7 shows the concentric pattern of rings, called "Newton's rings," resulting from a spherical surface in contact with a flat surface.

In Figure 8 a cylindrical surface rests on a plane surface. The cylinder is flat in one direction to produce straight fringes, but is curved in the other direction to produce a variable fringe width and spacing.

Figure 9 shows the interference pattern produced by an irregular glass surface, such as a microscope slide.

The procedure explained in the laboratory portion of this module is a standard test procedure used to test surface flatness of optical and machined surfaces. The interference of light is also used by a class of devices called "interferometers" to make measurements on optical components and systems. All interferometers function by splitting a light beam into two or more components, which travel along different paths and are recombined, forming an interference pattern. One such device, termed the "Michaelson interferometer," is displayed in Figure 10. This device is used to measure variations in thickness of an optical component by transmitted light.

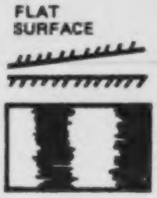


FIG. 5

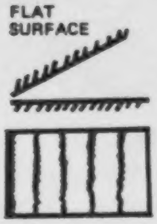


FIG. 6

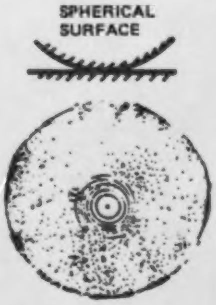


FIG. 7

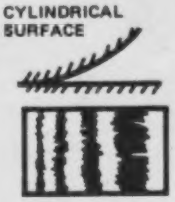


FIG. 8

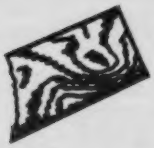


FIG. 9

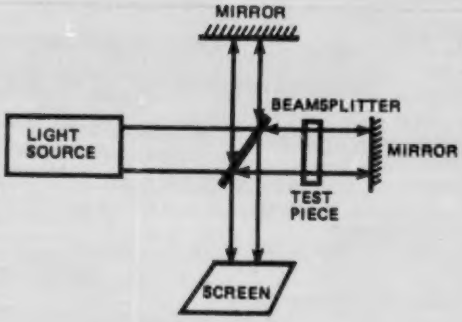


FIG. 10

SUMMARY

The interference of light waves makes possible extremely sensitive methods of testing the optical flatness of a surface (flatness compared with the wavelength of optical radiation, or visible light). The alternating light and dark fringes of interference patterns provide useful information about the surface dimensions of an object. Constructive interference produces bright lines; destructive interference, dark lines. If both the surface tested and the "master" are exactly flat (when the air wedge system is used), the pattern will consist of straight, evenly-spaced fringes.

AIR-WEDGE INTERFERENCE PATTERNS

MATERIALS

Mercury lamp with stand.
Dark green plastic filter.
1-foot square of black felt.
4" x 5" piece of ground glass with mount.
2-inch diameter optical flat, flatness of λ/10 (Edmund #1914 or equivalent).
Piece of flat glass.
Positive lens, focal length 1 m or greater.
Piece of plate glass.
Microscope slide.
Metric scale.

PROCEDURE

1. Set up equipment as shown in Figure 11. Be sure that all optical surfaces are clean and free of dust.
2. Look at flat glass from an angle to allow reflections of ground glass surface to enter eye. Focus eye on air wedge. Move flat glass until fringes are visible. Observe interference pattern; sketch in Data Table.
3. Repeat above procedure for positive lens, plate glass, and microscope slide.
4. Return flat glass to optical flat; obtain a fringe pattern.
5. Measure fringe spacing.
6. Complete Data Table.

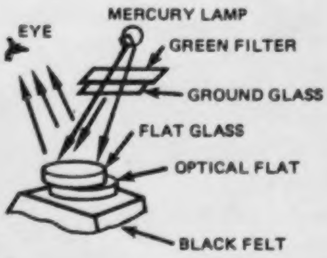


FIG. 11

DATA TABLE

Sketch the interference pattern produced by each surface; state what it reveals about that surface.

Flat glass:

Positive lens:

Plate glass:

Microscope slide:

CALCULATION OF WEDGE ANGLE (Flat glass on optical flat):

Fringe spacing: x = _____ cm = _____ m

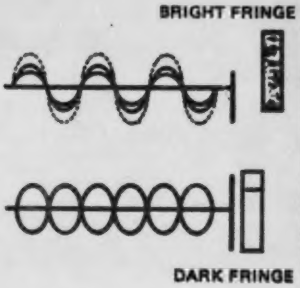
Wavelength of light: λ = 546 nm = _____ m

Wedge angle: $A = \frac{\lambda}{2x}$

$= \frac{\_\_\_\_\text{ m}}{2(\_\_\text{ m})}$

A = _____ rad

SELF CHECK

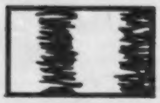


1. Indentify the adjoining sketches as constructive or destructive interference.

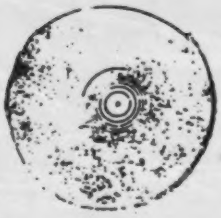
2. Draw and identify a diagram showing constructive and destructive interference of light reflected from air wedge between two glass surfaces.

3. Light from a helium-neon laser ($\lambda = 633$ nm) is used to illuminate an air wedge. Fringe spacing is 0.025 cm. Calculate the wedge angle.

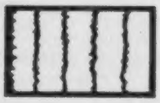
4. Identify the type of surface that would produce each of the following interference patterns if placed in contact with a flat surface:


(a)

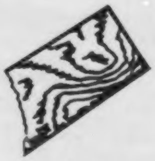
____


(b)

____


(c)

____


(d)

____

5. What is the wavelength of light that will produce a fringe spacing of 1 cm when the wedge angle is 4×10^{-6} radians?

6. An air wedge has an angle of 3×10^{-5} radians. What will the fringe spacing of the interference pattern be if the wavelength of the light producing the interference pattern is 700 nm?

13L12

DIFFRACTION GRATING — MEASURING THE WAVELENGTH OF LIGHT

Modern scientific investigation depends greatly on the analysis of light for its data. One of the most important characteristics of light is its wavelength. A diffraction grating is an optical device that separates light into its component wavelengths for application or measurement. This module discusses the phenomenon of diffraction of light and its application in diffraction gratings; a diffraction grating spectroscope will be used to measure the wavelength of light in the laboratory.

OBJECTIVES

To complete this module, the student should be able to:

- Summarize Huygens' principle in one or two sentences.
- Draw and label a diagram showing diffraction of waves at an edge barrier, including the edge barrier, shadow and wavefronts.
- Draw and label a diagram showing the intensity pattern resulting from double-slit diffraction, including intensity, orders of diffraction, distance, single-slit diffraction envelope and double-slit diffraction peaks.
- Given any two of the following quantities, calculate the third:

  Wavelength of light.
  Diffraction angle of a specific order.
  Grating spacing.
- Given appropriate equipment, measure the wavelength of light with a grating spectroscope.

DISCUSSION

Huygens' principle is used to predict the shape and position of a defined wavefront of light after it has traveled a given distance. It states that each point on a wavefront may be considered to be the source of a spherical wavelet. The new shape and position of the wavefront at any time is a line tangent to all the wavelets at that instant as shown for a spherical wavefront in Figure 1.

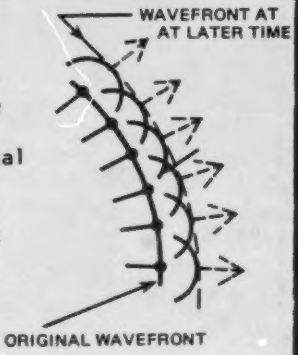


FIG. 1

Figure 2 shows this principle applied to a plane wavefront striking an edge barrier. Part of the wave bends around the edge of the barrier, then it travels into the area of the geometric shadow. This action is most easily observed for water waves, but is also true of all other waves, including light. This bending of a wave around an edge is called diffraction.

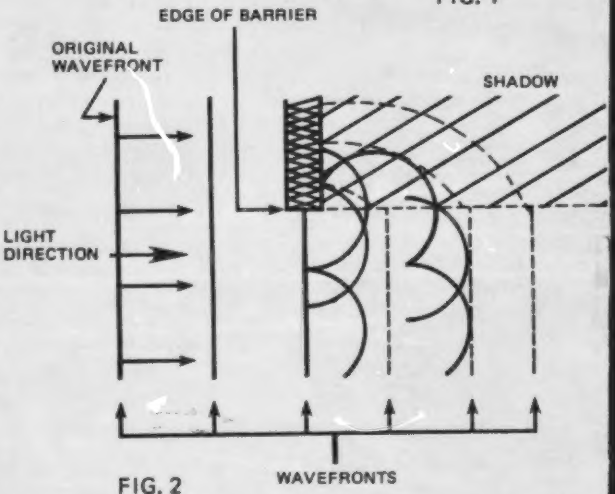


FIG. 2

Figure 3 shows plane wavefronts of light striking a narrow slit. The distance x to the screen is much greater than slit width b. The slit has been divided into several small segments; each segment is the source of a Huygens' wavelet. At point D, the wavelets from all slit segments are in phase, and constructive interference occurs to produce a bright area. At point E, the waves from A and C are in phase, but both are out of phase with other waves, for example, from B. The resulting destructive interference of all waves at point E produces a dark area. The intensity pattern produced on the screen (also shown in Figure 3) is called a single-slit diffraction pattern.

Figure 4 shows light waves diffracted through two slits. The arrows indicate lines along which the waves from the two slits are in phase (wavefronts intersect). Figure 5 is a diagram of double-slit diffraction. For point P to be a bright fringe, waves from the two slits must arrive in phase. For this event to occur, the difference in

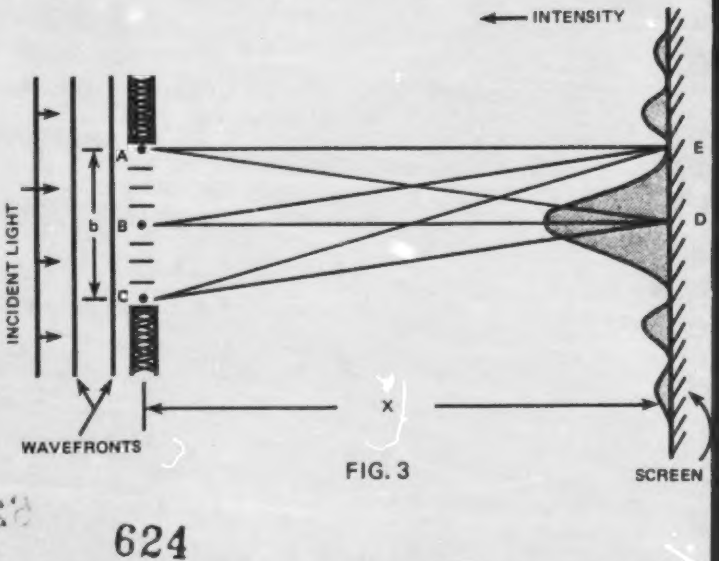


FIG. 3

optical path length for the two waves must be equal to an integral multiple of the wavelength of light. This optical path length difference is equal to d sin θ in Figure 5, where d is the slit separation and θ is the angle of diffraction. The equation for bright fringe formation is then:

$$d \sin \theta = m\lambda \qquad (1)$$

where

d = slit separation.
θ = angle of diffraction.
λ = wavelength of light.
$m = 0, 1, 2, 3 \ldots$

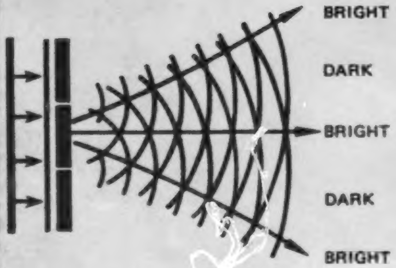


FIG. 4

This equation gives the angular location of the bright finges of a double-slit diffraction pattern. Figure 6 shows the intensity pattern resulting from double-slit diffraction. The envelope of the pattern (dashed line), which is the single-slit diffraction pattern, depends on the width of the slits. Each of the narrower peaks represents a point at which Equation 1 is satisfied for constructive interference between the two slits, and depends on their separation. Numbers in the circles are values of m (in the equation) for each peak, called "orders of diffraction." These numbers represent the whole number of wavelengths by which the optical path length of one slit differs from that of the other slit.

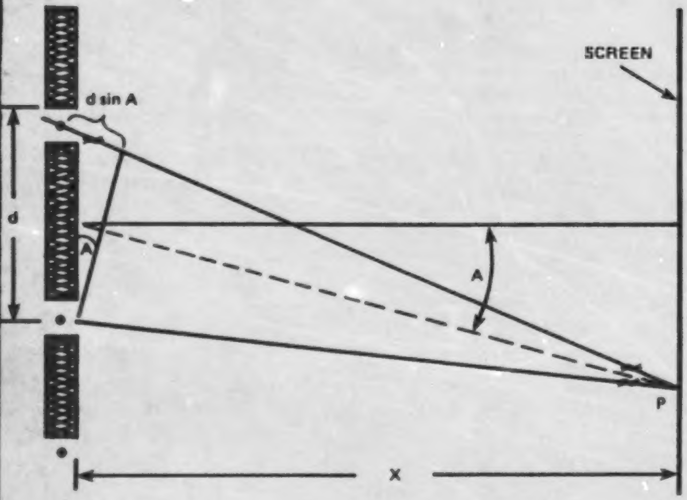


FIG. 5

Figure 7 shows a diffraction grating consisting of a large number of slits, each with a fixed width and spacing. Equation 1 also is true for the diffraction grating. Increasing the number of slits does not change the spacing of the diffraction pattern in Figure 6, as long as d remains fixed; however, increasing the

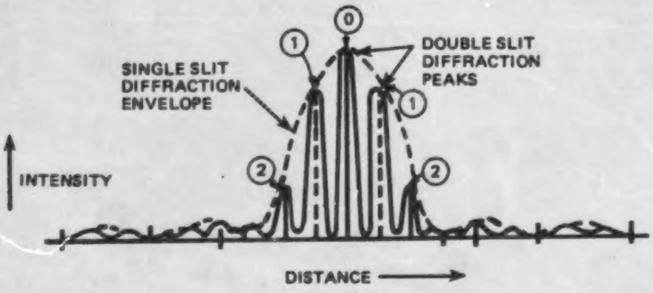


FIG. 6

number does decrease the width of peaks in the double-slit pattern, making separation of different wavelengths easier.

Figure 8 shows white light striking a diffraction grating with its slits perpendicular to the plane of the page. The zero order is undiffracted white light. On each side of this order are the higher orders of diffraction. In each order, red light is diffracted through a larger angle than blue light.

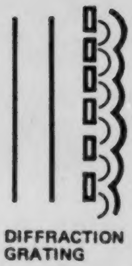

DIFFRACTION
GRATING

FIG. 7

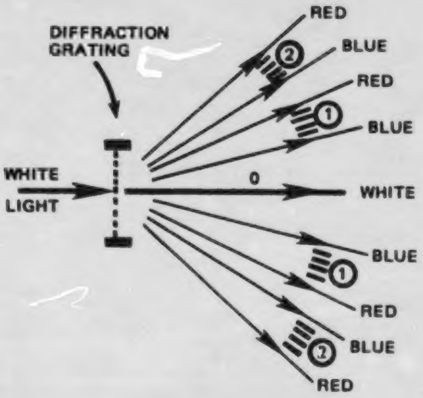

DIFFRACTION
GRATING
WHITE
LIGHT
0
WHITE
RED
②
BLUE
RED
①
BLUE
BLUE
①
RED
BLUE
②
RED

FIG. 8

Example A shows how data taken with a diffraction grating may be used to determine the wavelength of light by using Equation 1.

EXAMPLE A: CALCULATING WAVELENGTH OF LIGHT

Given: In Figure 8 the first order diffraction angle for red light from a helium-neon laser is 19.5°; the grating spacing is $d = 1.896 \times 10^{-6}$ m (13,400 lines/inch).

Find: Wavelength of laser light.

Solution: $d \sin \theta = m\lambda$

$$\lambda = \frac{d \sin \theta}{m}$$

$$= \frac{(1.896 \times 10^{-6} \text{ m})}{1} \times \sin (19.5°)$$

$$= (1.896 \times 10^{-6} \text{ m}) \times (0.3338)$$

$$= 6.33 \times 10^{-7} \text{ m}$$

$$\lambda = 633 \text{ nm.}$$

Examples B and C shown the solution to additional problems using this equation.

EXAMPLE B: DIFFRACTION ANGLE CALCULATION

Given: Grating spacing $d = 1.896 \times 10^{-6}$ m; wavelength of light $\lambda = 488$ nm (blue).

Find: First order diffraction angle.

Solution:

$$d \sin \theta = m\lambda$$

$$\sin \theta = \frac{m\lambda}{d}$$

$$= \frac{(1) \times (4.88 \times 10^{-7} \text{ m})}{(1.896 \times 10^{-6} \text{ m})}$$

$$= 0.2574$$

$$\theta = \sin^{-1} (0.2574)$$

$$\theta = 14.9°.$$

EXAMPLE C: CALCULATION OF GRATING SPACING

Given: Second order diffraction angle A = 37.72°; wavelength of light $\lambda = 580$ nm.

Find: Grating spacing.

Solution:

$$d \sin \theta = m\lambda$$

$$d = \frac{m\lambda}{\sin \theta}$$

$$= \frac{(2) \times (5.80 \times 10^{-7} \text{ m})}{\sin (37.72°)}$$

$$= \frac{1.16 \times 10^{-6} \text{ m}}{0.6118}$$

$$d = 1.896 \times 10^{-6} \text{ m.}$$

Figure 9 shows a reflection grating. Instead of a series of slits, this type of grating is composed of a series of reflective surfaces. Reflection gratings have much higher efficiency than transmission gratings because little light is absorbed. Angle B of the reflecting surfaces, with respect to the grating surface, is chosen to reflect the maximum amount of light in a particular direction. This angle is called the "blaze angle" of the grating.

Research grade gratings are designed to give the best diffraction efficiency for a particular wavelength of light in a particular order of diffraction. A number of optical instruments--including spectroscopes and monochromators--employ a diffraction grating as a principal element.

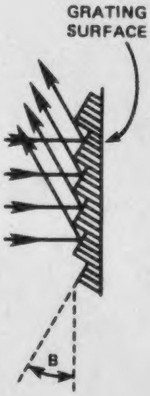


FIG. 9

SUMMARY

Diffraction is the bending of light around the edge of obstacle, causing a portion of the light energy to appear in the region of the geometric shadow. When light passes through two closely spaced, narrow slits, the diffracted light forms two sets of spherical wavefronts that interfere with one another and produce a pattern of light and dark on a screen. Location of the areas of maximum brightness depends on the separation of the slits and the wavelength of light. A diffraction grating, which consists of a large number of slits or reflecting planes, may be used to make highly accurate measurements of the wavelength of light.

MEASURING THE WAVELENGTH OF LIGHT WITH A DIFFRACTION GRATING

MATERIALS

Grating spectrometer (Sargent-Welch #S-75903-80 or equivalent) consisting of the following components:

Collimator with slit and lens that allows a small amount of light to strike grating in parallel rays.
Transmission-type diffraction grating with known spacing.
Telescope with objective lens, eyepiece and cross hairs.
Vernier scale to allow measurement of angles to nearest tenth of a degree.

Helium spectrum tube.
Sprectrum tube power supply.

PROCEDURE

1. Record grating spacing in Data Table.
2. Set up equipment as shown in Figure 10, but do not place grating on stage.
3. Turn spectrum tube power supply on; place spectrometer slit near capillary portion of tube. Beware of high voltage.
4. Align telescope with collimator; focus telescope to obtain sharp image of slit; set cross hair on image of slit.
5. Loosen set screw; move degree scale around until zero on degree scale lines up with zero on vernier scale; be sure cross hair is on image of slit; tighten set screw.

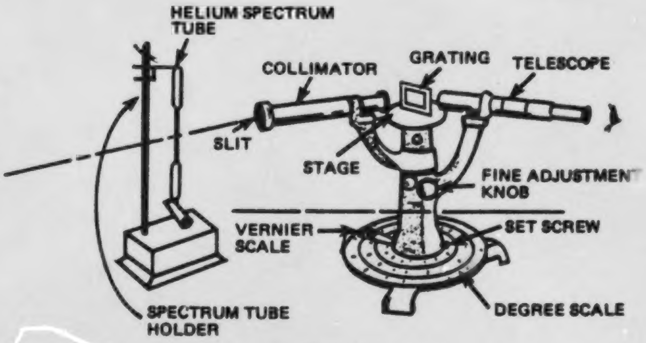

HELIUM SPECTRUM TUBE
COLLIMATOR
GRATING
TELESCOPE
SLIT
STAGE
FINE ADJUSTMEN
KNOB
VERNIER SCALE
SET SCREW
DEGREE SCALE
SPECTRUM TUBE HOLDER

FIG. 10

6. Place grating on stage, as shown in Figure 10, with rulings vertical and grating surface perpendicular to incident light; move telescope around until spectral lines (first order) are visible.

7. View spectral lines; refocus telescope; adjust slit width for sharpest image.

8. Set cross hairs on first violet line; read angle from vernier; record in Data Table.

9. Set cross hairs on a total of five strong lines in helium spectrum; record angle of diffraction for each.

10. Continue to move telescope until second order diffraction of spectral lines is visible; record diffraction angles for second order diffraction of same five lines.

11. Calculate wavelength of light for each angle measured; use grating equation.

12. Compare measured wavelengths to actual wavelengths given in Table I; state possible sources of errors in your measurement.

TABLE 1. STRONG HELIUM SPECTRAL LINES

Color	Wavelength (nm)	Color	Wavelength (nm)
Violet	388.6	Blue-green	492.2
Violet	402.6	Green	504.8
Violet	438.8	Yellow	587.6
Deep blue	447.1	Red	667.8
Blue	471.3	Red	706.8

DATA TABLE			
Color Order	Diffraction Angle°	Wavelength (nm) Experimental	Wavelength (nm) From Table
1			
2			
1			
2			
1			
2			
1			
2			
1			
2			

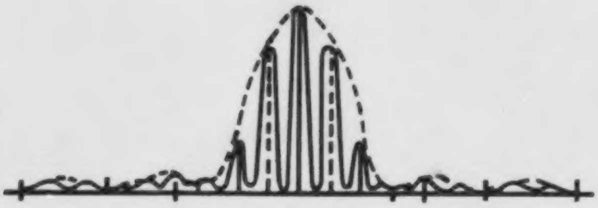
Grating spacing: d = _____m.

SELF CHECK

1. In one or two sentences, summarize Huygens' principle.

2. Sketch wave pattern produced when plane waves are diffracted at an edge barrier.

3. Label sketch (below); include several orders of diffraction, single-slit diffraction envelope and double-slit diffraction peaks.



4. A diffraction grating has a spacing of 2×10^{-6} m. Calculate diffraction angles for the following:

    First order of 400 nm light.
    First order of 650 nm light.
    Second order of 550 nm light.

5. A diffraction grating produces fourth-order diffraction of 633 nm at 42°. What is the grating spacing?

6. For the above grating, what wavelength has third-order diffraction at 42°?

7. A diffraction grating has 4000 lines/cm. At what angle will the first order bright fringes appear if the wavelength of the incident light is 650 nm?

8. A bright fringe occurs at an angle of 26°. If the diffraction grating has 2700 lines/cm and the wavelength of the incident light is 541 nm, what is the order of the fringe?

POLARIZATION — GLARE REDUCTION

Electromagnetic radiation is made up of time-varying electric and magnetic fields moving through space. The polarization of light is a description of the orientation of its electric field in space. This module discusses characteristics of polarized light and methods of producing and measuring polarization of light. The origin of glare and its reduction through the use of polarizing optical components is also presented. Polarization of laser light is measured in the laboratory.

OBJECTIVES

To complete this module, the student should be able to:

- Sketch the orientation of an electric field of unpolarized light, and of vertically and horizontally polarized light.
- Draw and label a diagram showing the orientation of an electric vector of light as it passes through a series of polarizers with transmission axes at specified angles.
- Calculate intensity of light transmitted by a polarizer, given the following:

  Maximum intensity transmitted when light is polarized parallel to transmission axis of polarizer.

  Angle between polarization direction and transmission axis of polarizer.

- Draw and label a diagram showing polarization by reflection.
- Explain in a short paragraph how polarized sunglasses reduce glare from horizontal surfaces.

- Given proper equipment, measure the transmission of a polizer/analyzer pair as a function of the angle between their transmission axes.
- Given proper equipment, measure Brewster's angle for a microscope slide in the laboratory.

DISCUSSION

Figure 1 shows an electromagnetic wave propagating through space. The electric field E, the magnetic field H and the direction of propagation x are all perpendicular to one another. Polarization of the electromagnetic wave describes the orientation of its electric vector in space. Thus, the wave in Figure 1 is plane polarized because 1) its electric field is always in one plane and 2) the plane does not change; that is, it is always vertical.

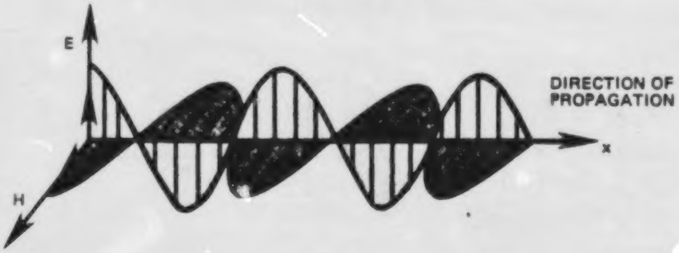


FIG. 1

There is no need to specify the direction of the magnetic field, because it always is perpendicular to both the electric field and the direction of propagation; magnetic field direction can be determined if these two facts are known.

Figure 2 shows three possible types of polarization and the symbol used to represent each in a diagram. In each case the figure on the right represents a side view and the figure on the left represents an end view of the ray of light. The electric field (Figure 2a) of the light is vertical (indicated by small vertical arrows on the light ray). Dots on the ray (Figure 2b) indicate horizonatal polarization; that is, it vibrates in and out of the page. Light that has random orientations of electric fields (Figure 2c) is called unpolarized light. This kind of light may be considered to be a combination of equal vertical and horizontal polarized components (indicated both by arrows and by dots).

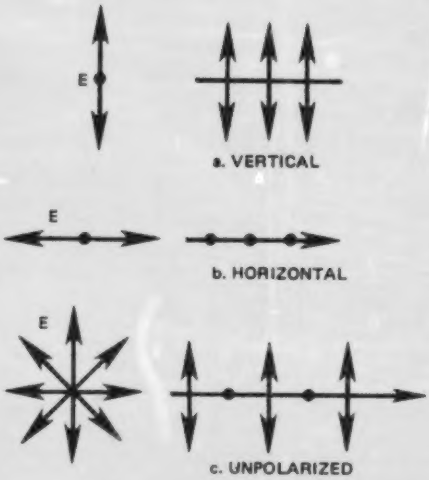


FIG. 2

A polarizer is a device that transmits only light of a particular polarization. Only that light with its electric field oriented in the same plane as the polarization axis is transmitted.

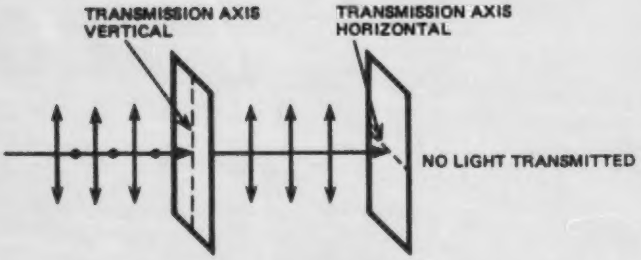

TRANSMISSION AXIS VERTICAL
TRANSMISSION AXIS HORIZONTAL
NO LIGHT TRANSMITTED

FIG. 3

In Figure 3, which shows unpolarized light striking a pair of polarizers, the first polarizer transmits the vertical component only. The second polarizer is oriented to transmit horizontally polarized light only; therefore, none of the vertically polarized light striking it is transmitted.

In Figure 4, which shows polarized light striking a polarizer with its plane of polarization making an angle A with the transmission axis of the material, the electric field E of the incident light can be resolved into two vector components: the part parallel to polarization axis E_y is transmitted; the component perpendicular to axis E_x is not. The transmission (ratio of light intensity transmitted) of polarized light through a polarizing material is given by Equation 1.

$$T = \cos^2 A. \qquad (1)$$

where
T = Transmission.
A = Angle between plane of polarization of light and transmission axis of polarizer.

Example A shows how to use this equation in solving a problem.

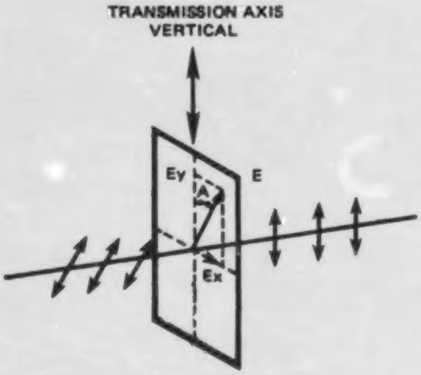

TRANSMISSION AXIS VERTICAL
Ey
A
E
Ex

FIG. 4

EXAMPLE A: TRANSMISSION THROUGH A POLARIZER

Given: Vertically polarized light strikes a polarizer that can be rotated to any orientation.

Find: Transmission, with transmission axis of polarizer oriented:

a. Vertical ($A = 0$).
b. At 45° to vertical ($A = 45°$).
c. Horizontal ($A = 90°$).

Solution: $T = \cos^2 A$

a. $T = \cos^2(0)$
$= (1)^2$
$T = 1.0$ (all light transmitted).

b. $T = \cos^2 (45°)$
$= (0.707)^2$
$T = 0.5$ (half of light transmitted).

c. $T = \cos^2 (90°)$
$T = 0$ (no light transmitted).

Figure 5 is a graph of transmission-versus-angle (between plane of polarization of light and axis of polarizer - angle A of Figure 4). As this curve indicates, transmission of a pair of polarizers can be controlled by rotating one of the polarizers. Some welding goggles incorporate a fixed polarizer and a second polarizer that may be rotated to adjust transmission.

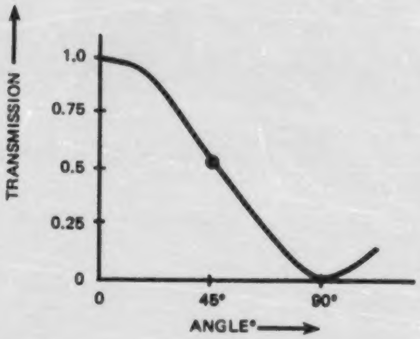


FIG. 5

Light may be polarized by several methods. The most familiar probably is polarization by absorption (Figure 6), which is used in polarizing sunglasses. The polarizer in Figure 6 consists of a plastic sheet containing needle-shaped crystals of herapathite (an iodine compound). These crystals absorb the horizontal components shown in the figure, but transmit most of the vertical components.

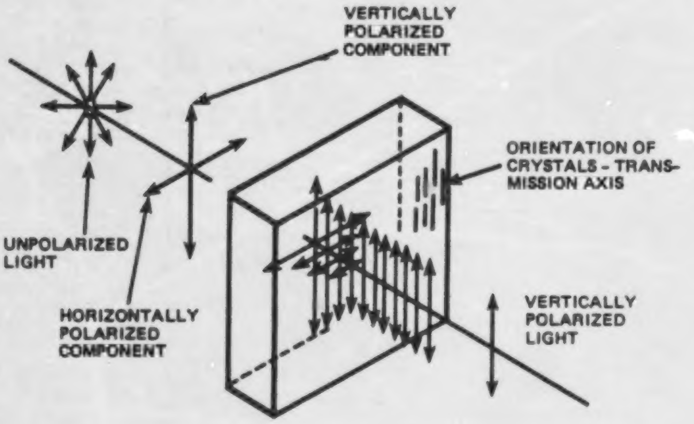

VERTICALLY POLARIZED COMPONENT
ORIENTATION OF CRYSTALS - TRANSMISSION AXIS
UNPOLARIZED LIGHT
HORIZONTALLY POLARIZED COMPONENT
VERTICALLY POLARIZED LIGHT

FIG. 6

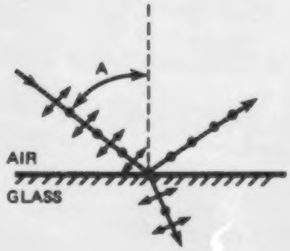


FIG. 7

Polarization by reflection of light from a surface is shown in Figure 7. Unpolarized light strikes surface with an angle of incidence A. A portion of the light is reflected, and the remainder is transmitted.

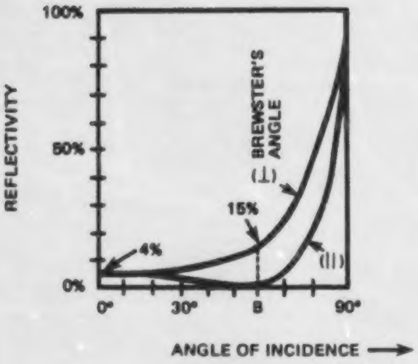


FIG. 8

The amount of each component (vertical and horizontal) reflected from the surface depends upon angle of incidence A (Figure 8). This graph indicates reflectivity of a glass surface as a function of 1) angle of incidence for light polarized in a plane perpendicular to surface (⊥) and 2) light polarized with its electric field parallel to the surface (||). Perpendicular components always reflect more strongly, except in the case of normal incidence; thus, light reflected from most surfaces tends to be at least partially polarized.

Angle B (Figure 8), which is called Brewster's angle, is the angle of incidence for which none of the parallel component is reflected. In Figure 7, angle A is Brewster's angle; its reflected ray is horizontally polarized.

The transmitted ray contains all of the vertically polarized component and about 85 percent of the original horizontally polarized component.

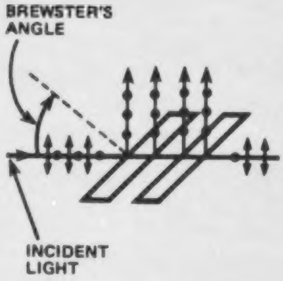


FIG. 9

Figure 9 shows a portion of a stacked-plate polarizer, which uses polarization by reflection to produce polarized light. Such a polarizer usually consists of four or more flat pieces of glass mounted at Brewster's angle. Reflected light is horizontally polarized, but the transmitted ray is not completely vertically polarized - it always contains some horizonatally polarized light not removed by reflection.

Brewster's angle for a material may be calculated by using Equation 2:

$$B = \tan^{-1} n$$

or

$$n = \tan B \qquad (2)$$

where
B = Brewster's angle.
n = index of refraction.

Example B illustrates the usefulness of this equation in solving a problem.

EXAMPLE B: BREWSTER'S ANGLE

Given: Index of refraction of glass: $n = 1.6$.

Find: Brewster's angle.

Solution:

$$B = \tan^{-1} n$$
$$= \tan^{-1} (1.6)$$
$$B = 58°$$

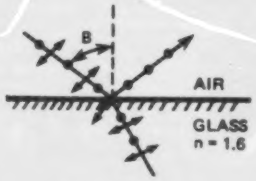


Figure 10 shows how the polarizing of sunglasses reduces glare. Most of the sunlight reflected from the surface of the water is horizontally polarized. Without the sunglasses this reflected glare is so bright that objects below the water surface cannot be seen. The lens material of the glasses absorbs all of the horizontal component and allows the vertical component to

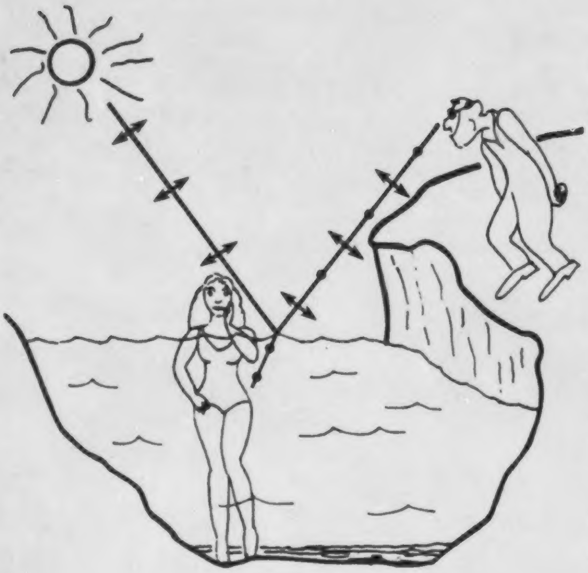


FIG. 10

pass through. The horizontally polarized glare is filtered out, and objects below the water surface are seen by vertically polarized light. This same principle applies to reflections from all horizontal surfaces.

The polarization axis of polarizing sunglasses is oriented vertically because most glare-producing surfaces (water, road surfaces, auto hoods and so on) are horizontal and produce horizontally polarized glare.

SUMMARY

Polarization refers to the orientation of the electric field of electromagnetic radiation in space. In unpolarized light, the electric field direction changes at random. In plane polarized light the electric field oscillates in a specific plane. Polarizers are optical components that transmit light of one polarization only. This transmission may be accomplished by reflecting light of other polarizations, as in a stacked-plate polarizer, or by absorbing unwanted light, as in the polarizing of sunglasses.

TRANSMISSION OF A POLARIZER

MATERIALS

Low-power unpolarized helium-neon laser (2 mW or less).
Optical power meter.
Optical bench.
Two lens holders.
Three carriages.
Two linear polarizers with axis of polarization indicated.
Lab jack.

PROCEDURE

1. Set up equipment as shown in Figure 11.

2. Turn laser on; observe all safety precautions.

3. Rotate polarizer nearest laser so that its transmission axis is vertical; this position yields vertically polarized light.

4. Rotate transmission axis of second polarizer to vertical.

5. Turn power meter on; carefully set dials to measure power of laser beam.

6. Make slight adjustments in orientation of second polarizer to achieve maximum transmission; record power in data table. If laser output power fluctuates, make all power readings at maximum point of needle swing.

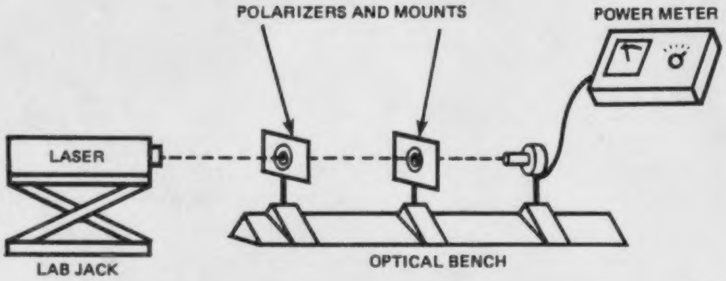


FIG. 11

7. Use pencil to make a small reference mark on polarizer mount just opposite zero degree mark on polarizer itself.

8. Rotate polarizer to bring 10° mark in line with reference mark.

9. Measure and record power.

10. Continue above process; take data every 10° until polarizer has been rotated 90°; change power meter scales as necessary; after each 90° make small adjustments to minimize transmission.

11. Calculate relative transmission (divide each power reading by the power at 0°).

12. Plot relative transmission-versus-angle on graph paper in data table.

13. Compare graph to that shown in Figure 5.

MEASURING BREWSTER'S ANGLE

MATERIALS

Low-power helium-neon laser.
Rotary table graduated in degrees, with stand.
Optical bench.
Two carriages.
Lens holder.
Polarizer.
Microscope slide.
Beeswax.
Index card.
Lab jack.

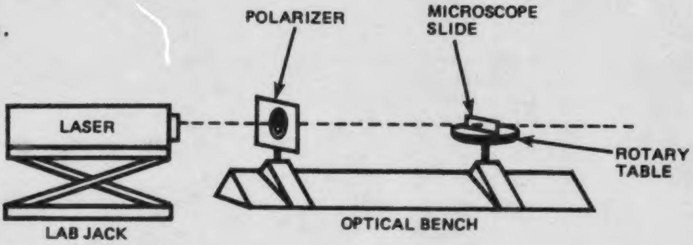


FIG. 12

PROCEDURE

1. Set up equipment as shown in Figure 12. Microscope slide is held in place by beeswax.

2. Turn laser on; observe all safety precautions.

3. Orient transmission axis of polarizer horizontally to produce horizontally polarized light - that is, vibrating in a plane parallel to top of rotary table.

4. Orient microscope slide to reflect laser beam back into laser output aperture; read and record angle on rotary table.

5. Slowly rotate table (with microscope slide) to position shown in Figure 13; observe reflected spot on index card.

6. Turn table to minimize intensity of reflected spot.

7. Measure and record new angular setting of rotary table.

8. Determine Brewster's angle (angle C in Figure 13); calculate index of refraction of slide (use Equation 2).

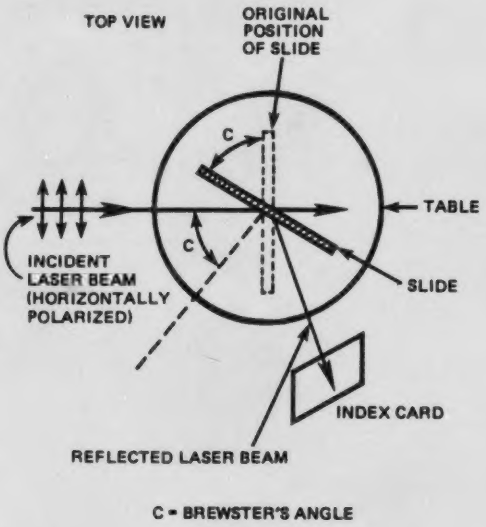
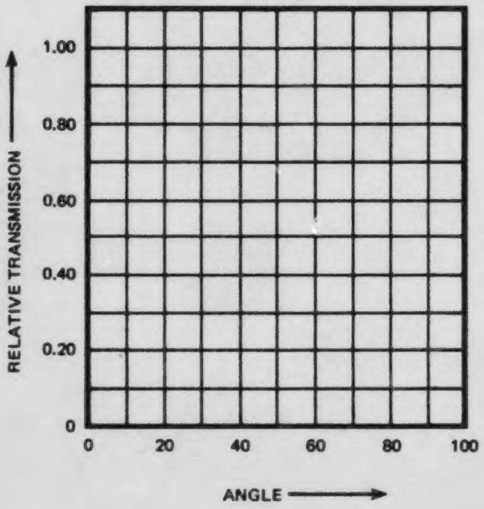

TOP VIEW
ORIGINAL POSITION OF SLIDE
C
C
TABLE
INCIDENT LASER BEAM (HORIZONTALLY POLARIZED)
SLIDE
INDEX CARD
REFLECTED LASER BEAM
C = BREWSTER'S ANGLE

FIG. 13



DATA TABLE		
Angle of Polarizer °	Power (mW)	Relative Transmission %
0		
10		
20		
30		
40		
50		
60		
70		
80		
90		

First angle measured: A = _____°.

Second angle measured: B = _____°.

Brewsters angle:

C = A - B = _____°.

Index of refraction: n = tan C

= tan (_____°).

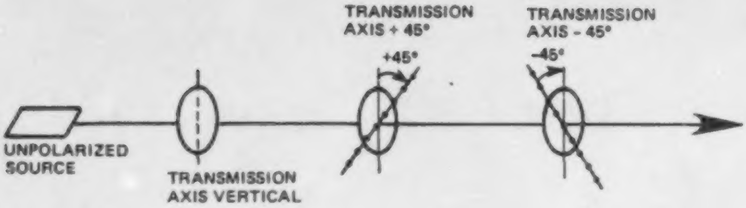
n = _____.

SELF CHECK

1. Sketch the electric fields of each of the following:

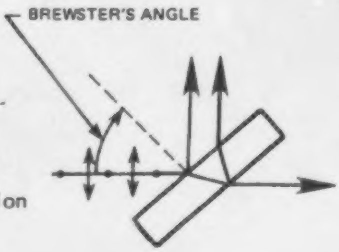
   Unpolarized light.
   Vertically polarized light.
   Horizontally polarized light.

2. Indicate (on diagram) the polarization of light as it travels through the following optical system:


TRANSMISSION AXIS + 45°
TRANSMISSION AXIS - 45°
+45°
-45°
UNPOLARIZED SOURCE
TRANSMISSION AXIS VERTICAL

3. If optical power through the first polarizer above is 10 mW, and there is no absorption and no reflection loss (perfect polarizer), calculate power transmitted by each of the other polarizers.

4. Unpolarized light strikes a piece of glass at Brewster's angle as shown on diagram (in the margin); indicate the polarization of 1) transmitted and 2) reflected rays.


BREWSTER'S ANGLE

5. Why is the axis of transmission vertical in polarized sunglasses?

6. For what angle between the plane of polarization of light and the transmission axis of the polarizer will transmission be a maximum?

7. For what angles between the plane of polarization of the light and the transmission axis of the polarizer will transmission be 0.365?

8. What is the index of refraction of a material for which Brewster's angle is 46°?

9. What is Brewster's angle for a material in which light travels at a speed of 2.4×10^8 m/s?

13L15

SPECTROSCOPY — LIGHT ANALYSIS

One of the most important analytical methods available to modern science is a process called spectroscopy. It is an analysis of light emitted by, reflected from, or transmitted through a material, and it may be used to determine a variety of physical characteristics of the material. This module discusses several types of optical instruments used in spectroscopy and presents a wide variety of applications of spectroscopy in science and industry. In the laboratory the student will use a spectrophotometer to determine the transmission curve of a filter.

OBJECTIVES

To complete this module, the student should be able to:

- Explain, in a brief paragraph, basic measurements performed in all types of spectroscopy.
- List three types of measurements that may be made by using spectroscopic techniques; give an application of each.
- Draw and label simple diagrams of following optical instruments:

  Spectrometer.
  Spectrograph.
  Monochromator.
  Spectrophotometer.

  Give an application of each.

- Explain, in a brief paragraph, how a spectrophotometer may be used to determine transmission of a sample for a particular wavelength of light.

- Given a spectrophotometer and a filter, draw transmission curve of filter for a wavelength (range of 350 nm to 700 nm).

DISCUSSION

TYPES OF SPECTROSCOPY

Spectroscopy is a process in which light of many wavelengths is dispersed to form a spectrum that may be observed or measured. In all types of spectroscopy the same two quantities must be determined: 1) wavelength of light present and 2) relative intensity at each wavelength. This information may be used to determine compositions and characteristics of materials producing, transmitting or reflecting the light.

Each type of atom or molecule has its own characteristic emission spectrum. The presence of a particular sequence of spectral lines in light emitted by a source indicates the presence of a particular atomic or molecular species in the source. Relative intensities at different wavelengths can reveal 1) relative amounts of each substance present and 2) source temperature. Study of emission spectra can lead to a determination of the composition of stars, flames and other light sources.

Each atom and molecule is capable of absorbing light with the same wavelengths as those of the light it can emit. If white light is passed through a partially transparent material, some of the light is absorbed. An analysis of the remaining (transmitted) light will give an absorption spectrum of the material. This technique can be used to analyze any substance that transmits light. It is a valuable tool in chemistry, physics, criminal investigation and medicine.

The analysis of light reflected from an object is also sometimes desirable. This technique is often used to check the quality and characteristic of mirrors and beamsplitters. Also, the reflected light can be used to reveal those wavelengths that have been absorbed by the material on reflection, and so functions in the same way as transmitted light.

INSTRUMENTS USED IN SPECTROSCOPY

A spectroscope is a device containing a prism or diffraction grating for dispersing light; it allows direct viewing of the sprectrum produced. If the device incorporates a method of measuring the wavelength of light, it also may be called a spectrometer.

Figure 1 shows a prism spectrometer that is commonly used in physics instruction. Such devices, though used primarily in education, employ the same principles as research instruments.

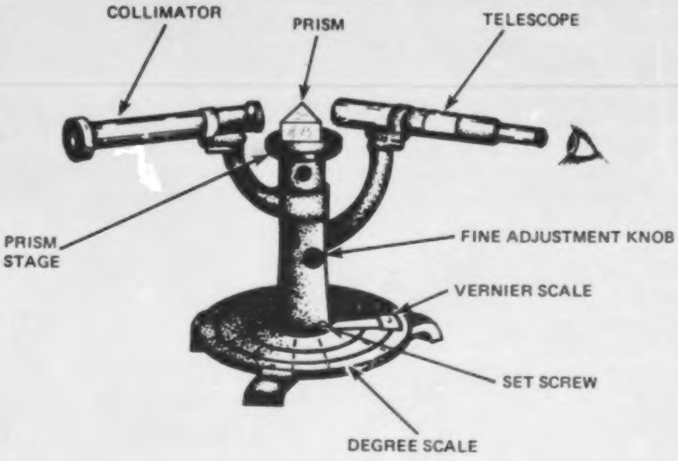


FIG. 1

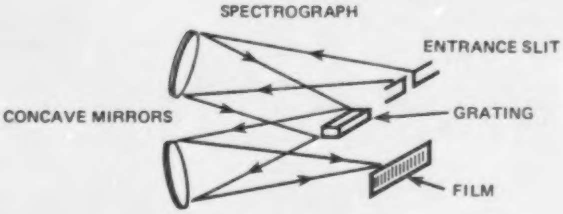


FIG. 2

Figure 2 shows a schematic of a grating spectrograph. This instrument is used to make a photographic record of the spectrum. Light enters through a narrow slit, is collimated by a concave mirror, then strikes a diffraction grating. The diffracted light is focused by a second concave mirror onto a photographic plate.

Figure 3 shows the emission spectra of several elements produced by this method. Each spectral "line" is an image of the entrance slit formed by a different wavelength present in the source light.

Spectrographs, may be used in the study of absorption spectra; however, they are used primarily to identify and measure wavelengths of lines emitted from or absorbed by known atoms and molecules

Figure 4 shows the optical components of a typical monochromator. This instrument is similar to a spectrograph but the photographic plate has been replaced with an exit slit. "White" light, containing a continuous disbution of wavelengths, is focused on the slit. Only a narrow range of wavelengths passes through the monochromator. The monochromatic wavelength emerging from the exit slit may be changed by rotating the grating. Monochromators are used in a variety of optical measurement systems to produce light of a specific desired wavelength; this process is used in measuring the absorption or reflection of materials. A monochromator, with one or more photomultipliers at its output slit, also is used in research to determine relative intensities of spectral lines.

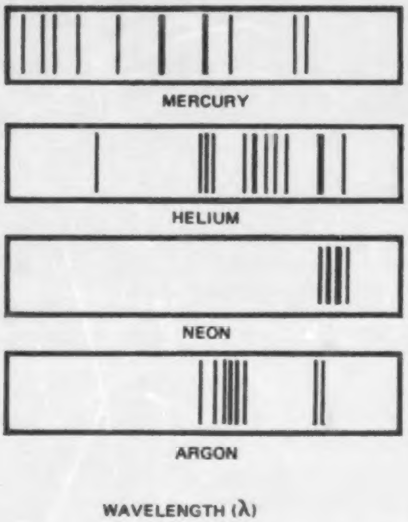

MERCURY
HELIUM
NEON
ARGON
WAVELENGTH (λ)

FIG. 3

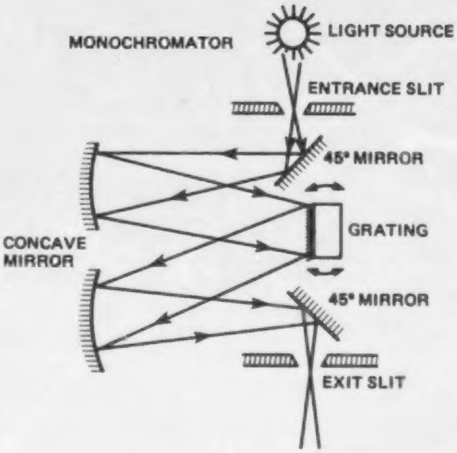

MONOCHROMATOR
LIGHT SOURCE
ENTRANCE SLIT
45° MIRROR
GRATING
CONCAVE MIRROR
45° MIRROR
EXIT SLIT

FIG. 4

Figure 5, a schematic diagram of a device called a spectrophotometer, shows that it consists of a white light source, a monochromator, a holder (cell, or curvette) for a sample of material to be analyzed, a photodetector and a display device. Such spectrophotometers are widely used as an analytical instruments in chemical laboratories. They also are used widely to measure reflectance and transmission of optical components. The more sophisticated instruments split the beam from the monochromator into two parts. One part goes directly to a detector while the other passes through the

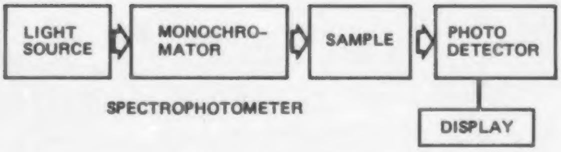

LIGHT SOURCE
MONOCHROMATOR
SAMPLE
PHOTO DETECTOR
DISPLAY
SPECTROPHOTOMETER

FIG. 5

sample to a detector. The two signal intensities are compared electronically and a permanent record of "transmission-versus-wavelength" is produced. Figure 6 shows a transmission curve for a piece of didymium filter glass, 4 mm thick, produced in this manner.

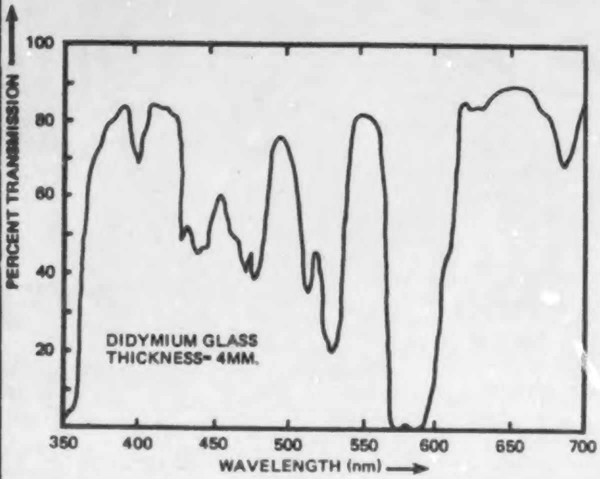


FIG. 6

Figure 7 shows a small, inexpensive spectrophotometer, sometimes called a colorimeter, commonly used in analytical laboratories. Figure 8 shows the optical components of this instrument. With no curvette in the sample holder, the exit slit is blocked by a shutter; thus, no light reaches the phototube, which allows the amplifier to be set at zero transmission with the amplifier control.

The wavelength control is used to select the wavelength of light allowed to pass through the exit slit of the monochromator section of the device. With reference curvette containing a clear liquid placed in the sample holder, the light control is adjusted to give a reading of 100 percent on the meter. The curvette is then replaced by an identical one containing the solution to be analyzed; the transmission is read from the meter.

The wavelength is changed and the process is continued until there are enough data points to plot a curve. To identify the unknown solution, this curve then may be compared with similar curves for known substances. The concentrations of dissolved substances also may be determined from this data.

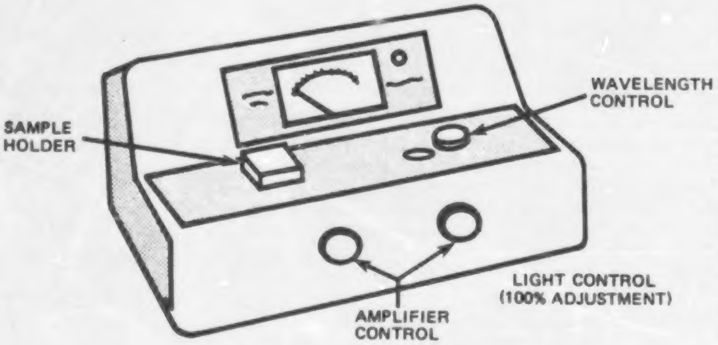
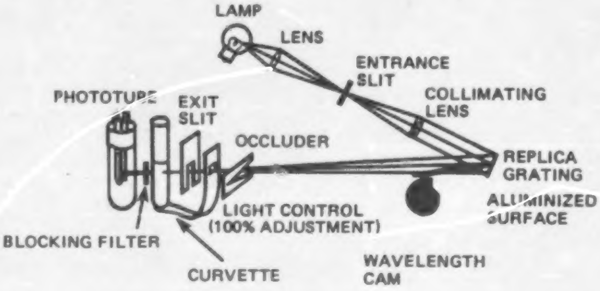


FIG. 7

SUMMARY

When light is emitted by, reflected from or transmitted through a material, the resulting wavelength distribution of the light energy carries information about the material. Spectroscopy is the name for a related group of optical measurement techniques that provide information about a material through analysis of the light with which it interacts. The two basic quantities determined in all types of spectroscopy are 1) the wavelengths of light present and 2) their relative intensities. Optical instruments used in these determinations include spectrometers, spectrographs, monochromators and spectrophotometers.


LAMP
LENS
ENTRANCE SLIT
COLLIMATING LENS
PHOTOTUBE
EXIT SLIT
OCCLUDER
REPLICA GRATING
ALUMINIZED SURFACE
LIGHT CONTROL (100% ADJUSTMENT)
BLOCKING FILTER
CURVETTE
WAVELENGTH CAM

LABORATORY

TRANSMISSION MEASUREMENT WITH A SPECTROPHOTOMETER

MATERIALS

Spectrophotometer
Didymium filter (Sargent Welch #S-75416-90).
Filter adaptor.
Curvette.
Distilled water.

PROCEDURE

1. Install filter adaptor in spectrophotometer.

2. Plug in and turn on spectrophotometer; adjust amplifier control knob (left front) until meter needle indicates 0% T; allow spectrophotometer to warm up for 15 minutes before adjusting.

3. Fill curvette to half full with distilled water.

4. Set wavelength control knob to 350 nm; re-zero needle if necessary.

5. Insert curvette filled with distilled water into filter adaptor.

6. Adjust light control (right front) for a meter reading of 100% transmission.

7. Insert didymium filter; read % transmission; record reading in data table.

8. Remove curvette and filter; set wavelength control to 375 nm; re-zero needle if necessary.

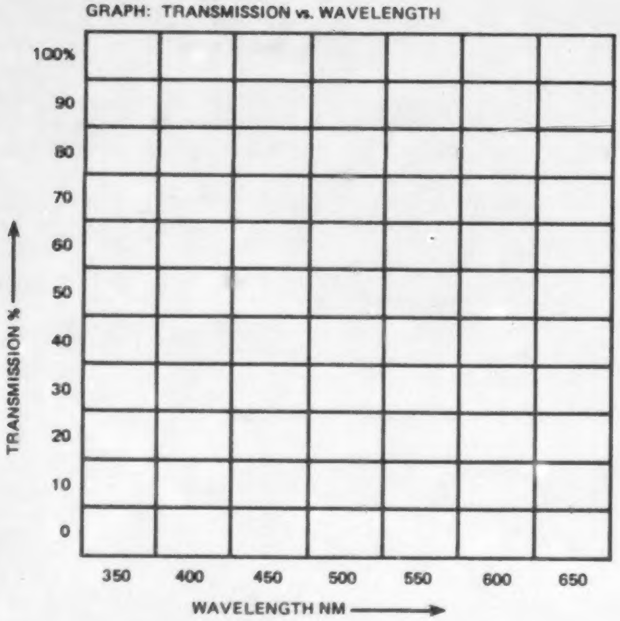
9. Repeat steps 5 through 7 for this wavelength.

10. Continue this procedure, increasing wavelength by about 25 nm each time; include measurements of minima and maxima and their wavelengths.

11. Draw a curve of "transmission-versus-wavelength" on the graph provided.

12. Compare your graph to Figure 6; explain differences, if any, in the two graphs.

DATA TABLE			
Wavelength (nm)	Transmission (%)	Wavelength (nm)	Transmission (%)



SELF CHECK

1. What two quantities must be determined in all types of spectroscopy?

2. Give an application of spectroscopic analysis for each of the following:

   a. Light emitted by a source.
   b. Light transmitted through a material.
   c. Light reflected from a surface.

3. Identify an optical instrument that has each of the following characteristics:

   a. Light of many wavelengths is diffracted by a grating onto a photographic film to make a permanent record of the spectrum.

b. Light of many wavelengths is diffracted by a grating or refracted by a prism to fall on a slit; light of only one color passes through the slit.

c. Light of many wavelengths is refracted by a prism; the resulting spectrum is viewed directly.

d. Light of a single wavelength passes through a material; the transmitted intensity is measured.

4. Identify the optical instrument used in each of the following applications:

   a. To display a visible sprectrum for viewing by physics students.

   b. To make a permanent record of an emission spectrum from a gas discharge for accurate wavelength determination.

   c. To obtain a light of a single wavelength for photocell calibration.

   d. To measure transmission of an unknown solution for chemical analysis.

5. List steps necessary for measuring (with a direct reading spectrophotometer) transmission of a filter at a particular wavelength.

6. For what wavelength is the percent transmission through a didymium glass, 4 mm thick, a maximum?

7. What frequency does 4 mm thick didymium glass transmit best?

8. What is the wavelength in glass (n = 1.50) of light that has a frequency of 5×10^{14} Hz?

9. A clear plastic material transmits only red light. What color will the plastic appear to be when white light is incident upon it?

10. What are the relative amplitudes of two light waves whose intensities are 40 mW and 60 mW?

RADIOACTIVITY

This module explores radioactivity, the emission of radiation during the decay of the unstable nuclei of atoms. Three different types of radiation are produced, denoted as "alpha," "beta," and "gamma." Radioactive atoms decay over a period of time and are characterized by a "half-life." For different radioisotopes, half-lives vary from a fraction of a second up to billions of years. Radioisotopes have important applications in science, in industry, and in medicine.

In the laboratory, the student will measure the decay rate and source strength of a radioactive source.

OBJECTIVES

Upon completion of this module, the student should be able to:

- Define the following terms:

  Radioactivity.
  Radioactive series.
  Half-life.
  Curie.

- State the relationship between the parent and daughter atoms for alpha-particle emission and for beta-particle emission.
- Given the necessary information about a radioactive source, calculate its half-life.

- Given the necessary information about radioactive properties of a sample of organic material, determine the sample's age.

- Given the necessary materials and equipment, calibrate a Geiger counter against a standard source and use the result to determine the decay rate, source strength, and number of atoms for a beta-ray emitter.

DISCUSSION

Radioactivity was discovered accidentally in 1896 by the French physicist Henri Becquerel. Becquerel placed some material containing uranium near photographic plates wrapped in black paper, and later found that the plates had been exposed. He correctly deduced that the uranium was emitting something that could penetrate the paper and expose the photographic plates.

Soon after, Pierre and Marie Curie discovered that ores containing uranium sometimes emit very intense radiation. They used chemical techniques to separate the active elements and, in so doing, identified two new chemical elements, radium and polonium. These two elements are more strongly radioactive than uranium. Madame Curie also correctly showed that the radiation was an atomic phenomenon. Over the next 20 years, studies by E. Rutherford and other workers led to the identification of the nature of the radiation produced by radioactive materials. In addition, many additional radioactive species were discovered. Today, many hundreds of radioactive species are known. Radioactive materials can be produced artificially, and such materials have found practical applications in science, in industry, and in medicine.

Radioactivity is defined as "the production of radiation in the spontaneous dacay of an atomic nucleus." Many types of atomic nuclei, some occurring in nature, are inherently unstable. They decay by themselves, giving off some type of radiation in the process. As a result of the decay, the atomic nucleus changes into a different type of nucleus.

RADIOISOTOPES

Radioactivity is exhibited only by certain isotopes. Recall that atoms with the same atomic number, but with different mass numbers, are called "isotopes." For a given chemical element (characterized by a certain atomic number), some isotopes may be radioactive while others are not. Thus, Pb^{210} is a radioactive isotope of lead; Pb^{208} is a stable isotope of lead and does not undergo radioactive decay. The isotopes that undergo radioactive decay are called "radioisotopes."

The early investigators of radioactivity identified three different types of radioactive emission. They arbitrarily named them "alpha," "beta," and "gamma" radiation. They could be separated by electric or magnetic fields. Figure 1 illustrates how a magnetic

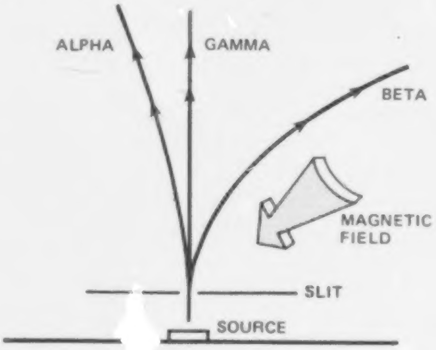


FIG. 1

field (perpendicular to the paper) affects the three types of particles. Alpha rays and beta rays are deflected in opposite directions, with beta rays being deflected more. Gamma rays are unaffected.

Today we know that alpha particles are the nuclei of helium atoms; that beta particles are electrons; and that gamma rays are electromagnetic radiation. Because the alpha and beta particles have opposite signs of electrical charge, they are deflected in opposite directions by electric and magnetic fields. Uncharged gamma rays are not deflected. For review, the properties of alpha, beta, and gamma radiation are summarized in Table 1. Notice that the alpha particle is both more massive and more highly charged than the beta particle.

TABLE 1. PROPERTIES OF ALPHA, BETA, AND GAMMA RADIATION

Type	Nature	Electrical Charge (Units of Electrical Charge)	Mass (gm)
Alpha	Helium nucleus	+2	6.62×10^{-24}
Beta	Electron	-1	9.1×10^{-28}
Gamma	Electromagnetic Radiation	0	0

A particular radioisotope decays in a definite way, either by alpha-particle emission or by beta-particle emission. All the alpha particles from a particular radioisotope have a definite value of energy, usually of the order of a few MeV. The energies of alpha particles from naturally-occurring radioisotopes are listed in handbooks for radiological health. Beta particles, however, have a spectrum of energies. The beta particle can have any value of energy, up to a maximum value characteristic of the particular radioisotope.

Now we consider the relation of the original nucleus, called the "parent," to the nucleus that remains after the decay, called the "daughter" nucleus. In alpha-particle emission, the emitted particle is a helium nucleus, a stable combination of two protons and two neutrons. Thus, in alpha-particle emission, the atomic number Z decreases by two, and the mass number A decreases by four. This decay is shown schematically in Figure 2.

The figure shows a gamma ray being emitted, in addition to the alpha particle. Alpha-particle emission frequently is accompanied by gamma radiation.

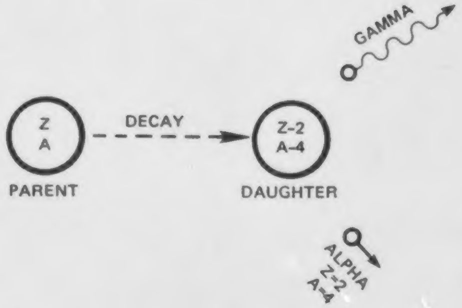


FIG. 2

In beta radiation, the emission is an electron, with one unit of negative charge (Figure 3). The atomic number Z of the nucleus increases by one, because in beta decay, a neutron in the nucleus converts to a proton, thereby increasing the positive charge while leaving the mass number unchanged.

Again, a gamma ray is shown. Usually, beta-particle emission is accompanied by gamma-ray emission. Beta emission always is accompanied by emission of another particle, called the "neutrino." The neutrino has no electrical charge, zero mass, and thus very weak interaction with material. It passes great distances through materials with almost no absorption.

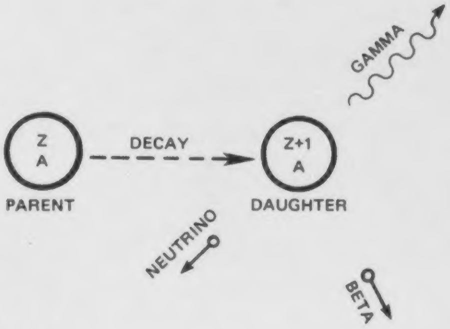


FIG. 3

Often, after a radioactive decay, the daughter nucleus also will be unstable and again undergo radioactive decay. The resulting nucleus again may be unstable. Thus, there can result chains of radioactive decay with a series of radioisotopes. The naturally-occurring radioisotopes at the upper end of the periodic table may be grouped into several different series. Table 2 lists the so-called "uranium series," which begins with U^{238} and ends with a stable isotope, Pb^{206}. The main chain of decay in the uranium series is shown; some of the relatively unimportant branchings of the chain are not shown.

TABLE 2. THE URANIUM SERIES (MAIN CHAIN)

ISOTOPE	ATOMIC NUMBER	RADIATION	HALF-LIFE
U^{238}	92	α	4.5×10^{9} years
Th^{234}	90	β	24 days
Pa^{234}	91	β	1.17 min
U^{234}	92	α	2.5×10^{5} years
Th^{230}	90	α	7.5×10^{4} years
Ra^{226}	88	α	1.62×10^{3} years

Rn^{222}	86	α	3.8 days
Po^{218}	84	α	3.05 min
Pb^{214}	82	β	27 min
Bi^{214}	83	β	20 min
Po^{214}	84	α	1.5×10^{-4} s
Pb^{210}	82	β	19.4 years
Bi^{210}	83	β	2.6×10^{6} years
Po^{210}	84	α	138 days
Pb^{206}	82	None	Stable

Inspection of the parent-daughter relationships in the chain is instructive in showing the changes in the nuclei produced by alpha or beta decay. In addition to the uranium series, there is the so-called "thorium series," which begins with Th^{232} and ends with stable Pb^{208}; and, the "actinium series," which begins with U^{235} and ends with stable Pb^{207}.

Most of the naturally-occurring radioisotopes are at the upper end of the periodic table, with atomic number of 83 or more, and are members of one of these series. There are a few naturally-occurring lighter radioisotopes, such as K^{40}. These decay to stable isotopes after only one or two steps and do not give rise to long series.

In addition to the radioactive species that occur in nature, there are hundreds of artificially-produced radioisotopes. The most common method of producing artificial radioisotopes is to bombard stable isotopes with neutrons in a nuclear reactor. A stable isotope (atomic number Z, mass number A) can capture a neutron and be changed to a new isotope with atomic number Z and mass number A + 1. In many cases, the new nucleus with the additional neutron will be unstable and will undergo radioactive decay. Radioisotopes of virtually all the naturally-occurring elements have been produced in this way. In addition, a number of elements with atomic number greater than 92 have been produced by bombardment of atoms with beams of atomic particles. Remember that uranium, with $Z = 93$, has the highest atomic number of any element found in nature. The so-called "transuranium" elements are extremely radioactive. Such elements have been produced with values of Z up to 106.

HALF-LIFE OF ELEMENTS

The half-life for a radioactive element is the time required for half the atoms to undergo radioactive decay. Thus, if we start with N atoms of a particular radioisotope, after one half-life has elapsed, there will be N/2 atoms remaining. After two half-lives, there will be N/4. After 3 half-lives, there will be N/8. Mathematically, the number of atoms remaining is expressed by Equation 1.

$$N(t) = N_o e^{-\left(\frac{0.693\ t}{T_{1/2}}\right)} \qquad (1)$$

where

$N(t)$ = number of atoms remaining after time t.
N_o = number at time zero.
$T_{1/2}$ = the half-life.

We see that when $t = T_{1/2}$, $N(t) = N_o e^{0.693} = N_o/2$, in accordance with the definition of half-life. Equation 1 is employed to solve a problem in Example A.

EXAMPLE A: NUMBER OF ATOMS REMAINING AFTER DECAY

Given: A certain radioisotope has a half-life of 10 days.

Find: The number of atoms remaining (a) after two weeks and (b) after four weeks.

Solution: (a) After 2 weeks, according to Equation 1 —

$$\frac{N\ (14\ \text{days})}{N_o} = e^{\left(\frac{0.693 \times 14}{10}\right)}$$
$$= 0.379$$

Thus, about 37.9% of the original atoms remain undecayed.

(b) After 4 weeks —

$$\frac{N\ (28\ \text{days})}{N_o} = e^{\left(\frac{0.693 \times 28}{10}\right)}$$
$$= 0.144$$

Therefore, 14.4% remain.

Notice that if one chooses the half-life $T_{1/2}$ in days, the time t also must be expressed in days. If $T_{1/2}$ is expressed in weeks, then t would also be expressed in weeks. All that matters is that t and $T_{1/2}$ be expressed in the same units of time.

The rate of decay for radioactive atoms is given by Equation 2.

$$Q(t) = \frac{0.693\ N(t)}{T_{1/2}} \qquad (2)$$

where

$Q(t)$ = rate of decay at time t; that is, the number of atoms decaying per unit time.
$N(t)$ = number of atoms present at time t.

Thus, as $N(t)$ decreases with time, the decay rate also decreases. The rate of decay for a source can be intense, according to Equation 2, either because the number of atoms is large, or because the half-life is short.

The half-lives of the radioisotopes in the uranium series of naturally-occurring radioisotopes have been listed in Table 2. The values cover a very broad range, varying from a minute fraction of a second up to billions of years. The naturally-occurring series are all headed by a long-lived radioisotope, such as U^{238}, which still remains from the original time of formation of the material of the earth. Other elements with atomic numbers greater than 92 may once have been present, but these elements had shorter half-lives and disappeared long ago. Short-lived isotopes in the uranium series are found in nature because they continually are being replenished by the decay of the isotopes above them in the chain. Thus, Pa^{234} is found in nature, despite its half-life of 1.17 minutes. It is being continually produced by the decay of Th^{234}, which in turn is formed from the decay of U^{238}.

The strength of a radioactive source is measured by the number of atoms that decay per second. The unit of source strength is the curie. A source with strength of 1 curie produces 3.7×10^{10} decays per second.

Equation 2 is clarified in Example B.

EXAMPLE B: DECAY RATE

Given: A source contains 1.5×10^{17} atoms of Ra^{226}.

Find: The activity of the source.

Solution: From Table 2, the half-life is 1.62×10^3 years, or 5.10×10^{10} seconds. From Equation 2, the decay rate is —

$$(0.693)(1.5 \times 10^{17})/5.1 \times 10^{10} = 2.04 \times 10^6 \text{ decays/s}$$

This value is equal to —

$$2.04 \times 10^6/3.7 \times 10^{10} = 5.51 \times 10^{-5} \text{ curies}$$
$$= 55.1 \text{ microcuires}$$

The number of particles that pass across a unit area in a unit of time from a radioactive source depends upon the distance from the source (assumed to be small, essentially a point) and upon the source strength expressed in numbers of decays per unit time. These quantities affect the intensity of radiation, as indicated by Equation 3.

$$I = \frac{S}{4\pi x^2} \qquad (3)$$

where

I = intensity of radiation (particles/cm^2-s).
S = source strength (measured as an activity, e.g., curies).
π = 3.14
x = distance from particles to source.

This expression neglects any absorption of the radiation. If the radiation is absorbed by the air, then the value given in Equation 3 will be reduced. The dependence in Equation 3 is commonly called an "inverse square law," because the intensity at a detector is inversely proportional to the square of the distance between the source and the detector. Equation 3 is used to solve a sample problem in Example C.

EXAMPLE C: RADIATION INTENSITY

Given: A P^{32} beta-ray source of strength 10 microcuries is in a vacuum, so that there is no absorption.

Find: The intensity of the beta radiation at a distance of 15 cm from the source.

Solution: The source emits $10 \times 3.7 \times 10^{10} \times 10^{-6} = 3.7 \times 10^5$ beta rays per second. According to Equation 3, the intensity is as follows:

$$I = \frac{S}{4\pi x^2}$$

$$= \frac{3.7 \times 10^5}{(4\pi)(15^2)}$$

I = 131 beta rays/cm^2-s

APPLICATIONS

Radioactive isotopes are used in many practical applications in medicine, in science, and in industry. In medicine, radioisotopes are used to test the functioning of certain organs of the body. For example, iodine taken into the body is collected in the thyroid gland. If a small amount of the radioisotope I^{131} is ingested, its location in the body can be followed with a counter. The intake of the I^{131} by the thyroid gland gives a measure of the functioning of the thyroid.

Radioisotopes also are used in medicine as therapeutic agents, often to destroy cancerous tumors. There are many irradiators that use the radioisotope Co^{60}. Gamma rays from Co^{60} are used to irradiate localized cancers, in an attempt to destroy them.

Radioactive decay of C^{14} is used to date biological samples. The isotope C^{14} is maintained in nature at a uniform level by cosmic rays* and is distributed throughout the world. Living organisms assimilate the carbon and incorporate it as part of their cells. The activity resulting from the decay of the carbon amounts to 15.3 disintegrations per minute per gram of carbon. When the organism dies, it stops incorporating C^{14}, and the C^{14} that it already has built into its cells gradually decays, with a half-life of 5568 years. If one measures the amount of activity from a biological sample and also determines the total amount of carbon by chemical means, one can obtain an estimate of the age of the sample, as illustrated in Example D.

EXAMPLE D: AGE OF A SAMPLE

Given: Someone shows you a fragment of charred wood that he claims came from the Trojan Horse (10,000 years of age). The weight of the wood is 100 gm. You measure the beta ray activity of the sample and find that it emits 1525 disintegrations per minute.

Find: Age of fragment.

Solution: The charred wood is mostly carbon. If it were 10,000 years old, the original number of C^{14} atoms should have decayed to $e^{(0.693 \times 10000/5568)}$ or 28.8% of the original value. The original number of decays was 15.3 per gram per minute, or 1530 from the total fragment. The current number of decays should be $0.288 \times 1530 = 441$ decays per minute. But the measured number (1525 decays/min) is much greater. Therefore, you decide that the wood is of much more modern origin.

---

*Assuming cosmic ray activity is constant.

Radioisotopes are used as heat sources in thermoelectric generators, where radiation is absorbed in the hot junction. The energy absorbed from the radiation keeps the junction hot. Generators powered by heat from radioactive decay have been used as sources of electrical power in remote locations and in the space program.

In industry, radioisotopes are used as gages to measure thickness of a material. The beta-ray source is on one side of the material and a counter on the other side. The counting rate depends upon the material thickness. Such gages can measure thickness of a moving material being produced in a continuous process, without any contact with the material. They are used to measure thickness of coatings applied to thin materials. They have been employed to measure thickness of products such as plastic sheet, paper, thin metal, rubber, and sandpaper. The output of the counter can be employed to control the production process and to ensure the proper thickness of the product.

Radioisotopes are capable of causing harm to human beings. Excessive exposure to radiation from radioisotopes can lead to many health problems, including genetic defects and cancer. Thus, strict safety regulations are applied to the transport, use, storage, and disposal of radioisotopes; on a national level, the Nuclear Regulatory Commission sets standards for their use. Many states also regulate the usage of radioisotopes. The user of radioisotopes must check the regulations relevant to the intended use and comply with them.

SUMMARY

Radioactivity is "the production of radiation in the spontaneous decay of an atomic nucleus." Isotopes (atoms having the same atomic numbers but different mass numbers) that undergo radioactive decay are termed "radioisotopes." The three types of radioactive emission are alpha, beta, and gamma radiation. Alpha particles (the nuclei of Helium atoms) possess a positive charge; beta particles (electrons) have a negative charge; and gamma rays (protons of high energy) have no charge.

The rate at which radioactive processes take place usually is measured by a quantity called the "half-life" of the material. Although a radioactive element loses energy, or "decays," at a rate, the whole period of its life cannot be calculated. Its half-life can be calculated, however, to determine the time required for half its atoms to change.

DECAY RATE AND SOURCE STRENGTH OF K^{40}

MATERIALS

Radioactive sources:
- U^{238} source — chemically purified and of known mass.
- K^{40} source.

Geiger counter, with thin window (<0.02 cm thick aluminum, or equivalent), with scalar and high-voltage power supply.
Aluminum-foil absorbers — 0.05 mm thick.
Holders for sources, counter, and aluminum foils.
Meter stick.
Semilogarithmic graph paper.

PROCEDURE

The student will measure the decay rate and source strength for K^{40}, a beta-particle emitter. The radioisotope K^{40} decays according to —

$$K^{40} \longrightarrow Ca^{40} + \beta + \nu,$$

where β represents the beta particle and ν the neutrino. The isotope Ca^{40} is stable. The maximum energy of the beta spectrum is 1.36 MeV. (This value neglects a relatively small competing mode of decay.)

In order to measure decay rate, one should determine the counting rate of a detector viewing the source and then determine the fraction of the particles counted by the counter (that is, the counter efficiency). The establishment of counter efficiency, unfortunately, is very difficult. In practice, a calibrated source is easier to use to determine detector response, and, hence, to obtain the decay rate of the unknown source.

For the calibration, we will use beta rays from Pa^{234}. The source will be a chemically-purified source of U^{238}. After a U^{238} source is prepared, it is free of atoms of the daughter products. Within a reasonably short time, there also will be decays of Th^{234} and Pa^{234}, by virtue of buildup of these isotopes following alpha decay of U^{238} atoms. (See Table 2.) Because of the relatively long half-life of U^{238}, there will be virtually no decays from radioisotopes lower in the series. Moreover, the beta rays from Th^{234} are very low-energy ones and are removed easily by an aluminum absorber with a thickness of 0.02 cm. Thus, one may have a situation in which the counts from the Geiger counter are dominated by beta rays from Pa^{234}.

If one has two sources, the unknown K^{40} source, and the U^{238} source of known mass, one can calibrate the detector, using the U^{238} source. Then one can use the result to determine the decay rate and the source strength of the unknown source. For the calibration, one uses the known result that Pa^{234} will decay at the rate of 1.22×10^4 beta disintegrations per second for each gram of U^{238} initially present.

1. Mount U^{238} source on sample holder. Mount Geiger counter on its holder, with the thin aluminum window facing toward source. Place source and sample close together, leaving just enough room to insert absorbing foils between them.

2. Measure counting rate from the source (mainly due to Pa^{234} beta particles). Insert aluminum foils one-by-one between source and counter, and measure counting rate after each foil is inserted. Continue until counting rate has dropped to less than 10% of its original value.

3. Plot counting rate versus absorber thickness on semilogarithmic graph paper. Remember to include thickness of counter window in absorber thickness. The curve should be approximately a straight line over a reasonable range of thickness.

4. Extrapolate the straight line portion of the curve back to zero thickness in order to obtain the counting rate at zero thickness. Call this number: $I_o\ (Pa^{234})$.

5. Determine the counter efficiency η as follows: The total number of disintegrations from Pa^{234} is $(1.22 \times 10^4) \times M$ disintegrations per second, where M is the mass of the U^{238} source in grams. The efficiency η is —

$$\eta = \frac{I_o\ (Pa^{238})}{(1.22 \times 10^4)M}$$

6. Replace the source with the unknown K^{40} source. Repeat steps 1-4, and obtain a counting rate for zero absorber thickness, $I_o\ (K^{40})$.

7. Obtain the total disintegration rate of K^{40} given by $I_o\ (K^{40})/\eta$. This is expressed in decays per second.

8. Calculate source strength in curies. (Recall, 1 curie = 3.7×10^{10} decays per second.)

9. Calculate number of K^{40} atoms in source. The half-life is 1.6×10^4 years. Use Equation 2.

SELF CHECK

1. Define the following terms:

   Radioactivity.
   Naturally-occurring chains of radioactive isotopes.
   Half-life of a radioisotope.
   Curie.

2. State the relationship between atomic number and mass number of the parent and daughter atoms for alpha-particle emission and for beta-particle emission.

3. Three days ago, you had 1.25×10^{19} atoms of a certain radioisotope. Because of radioactive decay, you now have 0.57×10^{18} atoms. What is the half-life of the isotope?

4. While you are digging at an archeological site, you find a fragment of old human bone. By chemical methods, you determine that the bone contains 225 gm of carbon. You measure the beta-ray activity of the material and find that there are 2684 C^{14} decays per minute. How old do you estimate the bone to be?

5. What is the source strength of a source for which 300 particles cross a 1 cm^2 area, located 0.1 meter from the source, each second.

6. A radioactive source has a half life of 1000 years. What percentage of the original number of atoms remain undecayed after 10,000 years?

7. After 6 days only one-tenth of the original number of atoms in a radioactive source remain undecayed. What is the half-life of this source?

APPENDIX

ADDITIONAL APPLICATION MODULES AVAILABLE FOR UTC PHYSICS

8M2, WEDGES AND INCLINED PLANES

A wedge is a device that transforms a small force applied in one direction into a pair of larger opposing forces almost perpendicular to the original force. This module describes the wedge as a pair of inclined planes and provides calculations of the mechanical advantage and efficiency of a wedge. In the laboratory, the student will use spring scales to determine the mechanical advantage and efficiency of a wedge.

8M5, BLOCK AND TACKLE

This module presents the block and tackle as a pulley system that operates on the principle of lever action. Various pulley systems are discussed, and examples are provided for the calculation of mechanical advantage and efficiency. In the laboratory, the student will use a spring balance and weights to determine the mechanical advantage and efficiency of several pulley systems.

8M6, DIFFERENTIAL HOIST

In this module, the wheel and axle are included in a block and tackle system in order to produce a differential hoist, or chain hoist. The differential hoist is described, and calculations for its mechanical advantage and efficiency are provided. In the laboratory, the student will use a spring balance and weights to determine the mechanical advantage and efficiency of a differential hoist.

9F3, WIND POWER

Windpower is produced by machines that convert wind energy into mechanical energy, which can be used for a variety of purposes. In this module, power delivered by a windmill is discussed, and various wind-power devices and their characteristics are presented. In the laboratory, the student will use a simple wind generator to produce electricity and will use an anemometer, a voltmeter, and an ammeter to determine the conversion efficiency of wind energy into electrical energy.

9T1, BOILERS

Boilers convert the thermal energy of burning fuel into the potential and kinetic energy of steam. This module discusses the energy conversion from hot combustion gases to pressurized steam in a boiler. In the laboratory, the student will determine the effects of preheated water and of ash deposits upon steam generation. A miniature steam engine is utilized.

10F2, ANEROID BAROMETERS

An aneroid barometer is an energy transformer that accepts input energy from the pressure of a gas surrounding it and transforms it into potential energy in a spring. A pointer, or other display device, indicates the amount of energy stored in the spring. In this module, the construction and applications of an aneroid barometer are investigated. In the laboratory, the student will measure a variety of pressures with an aneroid barometer and will express each pressure in several standard pressure units.

10T1, AUTOMOBILE THERMOSTATS

This module presents the automobile thermostat as a device that employs the principle of thermal expansion to control the temperature of the automobile engine. As the water in the engine block is heated, thermal expansion in the thermostatic device opens a valve that allows the water to flow to the radiator. In the laboratory, the student will examine the operations of an automobile thermostat.

11M1, SPEED OF SOUND IN MATERIALS

Sound waves travel through various materials with a speed dependent upon the elastic properties of the materials, characterized for each material by a constant called the "modulus of elasticity." This module discusses the velocity of sound in different types of materials. In the laboratory, the student will use an oscilloscope and a microphone to determine the velocity of sound in different media.

11M2, REFLECTION AND ABSORPTION OF SOUND ENERGY

This module discusses sound waves, echoes, and sound absorbent materials. Absorption and reflection coefficients, sound-level measurements, and exposure to noise are examined. In the laboratory, the student will use a sound level meter or a microphone and oscilloscope to determine sound levels transmitted through, and reflected from, various materials.

11M5, SOUND SPEAKERS

A sound speaker is a device that converts electrical energy into sound energy. In this module, the design and performance of several types of speakers is examined. Emphasis is placed on construction details of speakers that produce high fidelity sound over specific frequency ranges. In the laboratory, the student will use an audio signal generator, a microphone, an audio amplifier, and an oscilloscope to determine the frequency response of a moving coil speaker.

11F2, SPEAKER ENCLOSURES

This module discusses the characteristics of bass reflex, horn, and full speaker enclosures. In the laboratory, the student will measure directionality of sound produced by a speaker with and without an enclosure. Materials to be used include an audio signal generator, an oscilloscope, a speaker, a microphone, and an audio amplifier.

11F3, REVERBERATION TIME

This module describes the reverberation time of sound waves as part of the study of acoustics. The basic principles of acoustics are discussed, and an example of the calculation of reverberation time for a room is provided. In the laboratory, the student will measure the dimensions of the various surfaces in a room and will calculate the reverberation time of the room.

11F5, PRESSURE MEASUREMENT OF SOUND WAVES

Sound waves are small variations in pressure superimposed upon the normal atmospheric pressure. These small changes from atmospheric pressure are called "sound pressure." This module examines the elementary principles that characterize the production, transmission, and measurement of sound. In the laboratory, the student will use a sound level meter to measure sound-pressure levels and will draw the contours of equal sound-pressure levels around a sound source.

11E5, ELECTRONIC FILTERING IN d.c. POWER SUPPLIES

Electronic filters are used in d.c. power supplies to reduce the ripple of a.c. components at the load. Choice of the filter type to be used depends upon the ripple factor, regulation, and current of the power supply. This module discusses the construction, characteristics, and applications of several filter types. In the laboratory, the student will use electronic components, a multimeter, and an oscilloscope to construct and evaluate several filters.

11E6, ELECTRONIC FILTERING FOR FREQUENCY SELECTION

Electronic filtering for frequency selection is accomplished by means of reactive circuits. Reactive circuits are composed of capacitors and inductors that react to changes in the amplitude of a signal. Frequency dependence of capacitive and inductive reactance permits their use in filtering for frequency selection. This module examines high-pass, low-pass, selective, notch, band-pass, and band-stop filters. In the laboratory, the student will use a signal generator, an oscilloscope, and an ohmmeter to investigate the frequency selection of various filters.

11E7, MODULATION OF ELECTRICAL WAVES

Information is transmitted through space in the form of varying electromagnetic waves. The process by which information is introduced onto a transmitted wave is called "modulation." This module discusses amplitude modulation (AM) and frequency modulation (FM), which are the two basic types of modulation used in modern communications. In the laboratory, the student will use an RF signal generator, an audio signal generator, and an oscilloscope to observe amplitude modulation of waves.

12M1, SHOCK ISOLATION BY ABSORBERS

Shock isolation in a mechanical system is accomplished by absorbers that damp mechanical oscillations transmitted from one portion of the system to another. This module examines the principles and applications of vibration isolation systems. In the laboratory, the student will use a piezoelectric accelerometer and an oscilloscope to observe vibration isolation in simple mechanical systems.

12E1, RC CIRCUITS

In electronic circuits, it often is necessary to provide some kind of timing signal to synchronize events. One of the most common sources of such a signal is a resistance-capacitance (RC) circuit. This circuit generates regularly-spaced pulses by charging and discharging a capacitor through a resistor. This module examines the characteristics of RC timing circuits and the time constant. In the laboratory, the student will construct an RC timing circuit and test it with an oscilloscope.

12E2, RL CIRCUITS

An RL series circuit consists of resistance R in series with an inductor of self-inductance L. To distinguish between the effects of resistance and self-inductance, the RL circuit analyzed in this module consists of an ideal, resistanceless inductor in series with a separate noninductive resistor. An example of a real inductor with resistance is included. In the laboratory, the student will use an oscilloscope, a signal generator, and a VOM to determine the exponential behavior of current and voltage, and will measure the time constant associated with their decay.

13L14, HOLOGRAPHY - THREE DIMENSIONAL PHOTOGRAPHY

Holography is a method of using coherent laser light to record and reproduce three-dimensional images. This module discusses the differences between holography and ordinary photography, techniques used in making and viewing holographic images, and several applications of holography. In the laboratory, the student will use a helium-neon laser, a diverging lens, and a hologram to observe the three-dimensional image produced by holography.

13L16, INFRARED RADIATION THERMOMETERS

This module discusses the way in which the infrared radiation emitted from most objects can be used to determine the object's temperature. The design and limitations of a simple infrared radiation thermometer are presented. In the laboratory, the student will make practical measurements with an infrared radiation thermometer.

13L17, ILLUMINATION

This module examines different quantities associated with the measurement of radiant energy. The efficiency and spectral composition of various light sources are presented. The principles of operation and practical applications of light-measuring devices are examined, and minimum lighting requirements for various activities are discussed. In the laboratory, the student will use a photoelectric light meter to measure the illumination of several rooms.

13L18, SINGLE LENS REFLEX CAMERA

A single lens reflex camera (SLR) allows the photographer to view a scene, to focus, and to measure light intensity through the same lens used in taking photographs. This module introduces the optics, shutter speed, f-stop (lens opening), metering, film, viewing, and focusing of the SLR camera. In the laboratory, the student will use a convex lens, a Fresnel field lens, a pentaroof prism, and a Polariod camera to determine image characteristics and depth of focus.

13N1, GEIGER COUNTERS

This module presents the Geiger counter as an instrument used in the detection and measurement of radiation. The Geiger-Mueller detector is a gas-filled tube that detects radioactive particles, which ionize the atoms of the gas. This ionization produces electrical pulses that register the intensity of radiation. In the laboratory, the student will use an electronic scaler to determine the operating voltage of a Geiger counter, and will measure ionizing radiation produced by a radioactive source.

13N2, PENETRATING POWER OF ALPHA AND BETA RADIATION

Alpha and beta particles lose energy as they pass through a material because they collide with the electrons in the materials and give up energy to them. The distance the particle is able to travel in the material is its "range." This module describes the range of a particle as a function of its energy. In the laboratory, the student will use a Geiger counter to determine the range of alpha and beta particles in metal foils.

13N4, ABSORPTION OF GAMMA RADIATION

This module discusses the absorption of gamma radiation, which is an electromagnetic radiation that has a wavelength shorter than that of light. The absorption of gamma radiation is described by the same type of exponential equation that characterizes the absorption of light. The absorption coefficient for gamma radiation in a material depends upon the wavelength of the radiation and upon the properties of the absorbing material. In the laboratory, the student will use a Geiger counter to measure the transmission of gamma radiation through lead and aluminum, and will use the results to calculate absorption coefficients for these two materials.